Engineering Set-Theoretic Concepts

Neil Barton

13 August 2022
Words: Approx 32′000

Preface and acknowledgements

[NAB: Give acknowledgements here.]

Contents

Chapter 1

Introduction

If you're reading this book then I presume that you have some interest in 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 though 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 a very natural place to start. Understandably, philosophers have showed a lot of interest in set theory since its inception around the turn of the 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 an introductory logic course under their belt.

On the other side of the coin a great deal of work has appeared in our current century reflecting deeply on the nature of our concepts, and how we design, evaluate, and implement them. I'll follow much of the philosophical community in terming this discipline *conceptual engineering*. Whilst this is not a new idea in philosophy, it has certainly seen a greater focus in the last twenty years and has emerged as its own field of study. I think it's fair to say (again with a few notable exceptions[1]) that this literature has largely been neglected in the philosophy of set theory, and perhaps the philosophy of mathematics more widely. Again the issues in conceptual engineering are subtle, and many philosophers of set theory have been concerned with employing their finite time in the service of developing the relevant programmes, and so this lack of consideration is perfectly understandable. I do think, however, that viewing the philosophy of set theory in this conceptual light reveals some interesting possibilities for developing the field.

With these two aspects of the current philosophical landscape in mind, I have the following two aims in this short monograph:

(1.) I want 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.

(2.) I want to propose the idea that we are *now* at a conceptual crossroads ourselves, and that there are distinct concepts of set at play in our thought concerning sets.

Here's how I'll try to achieve these two aims:

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

[1] For example [Incurvati, 2020] and [Tanswell, 2018].

desiderata that will be employed later in the book when we come to assess set-theoretic concepts/conceptions. Chapter 3 will then survey some (a small relevant fragment) of the literature on *conceptual engineering*. We'll see a very important distinction for our purposes—that of the difference between a *concept* and a *conception*—and we'll talk about how these concepts/conceptions can be evaluated, engineered, and implemented. Chapter 4 will go over the naive conception of set and the paradoxes of that brought it down. We'll also provide a diagnosis of the problem as involving a conflict of two inconsistent constitutive principles for the naive conception. This material is well-worn, but I'll provide a twist on the classic paradoxes that has been noticed recently (that we can think of these paradoxes as paradoxes about the existence of *functions*) and this will help integrate this material with some more recent developments that will come later in the book. Chapter 5 will argue that we've *already* seen conceptual engineering in set theory, and we'll discuss the emergence of the *iterative* and *stratified* conceptions of set in the 20th century. We'll see that our intellectual ancestors faced difficult conceptual choices, and this will help us to see later that we ourselves are situated at a similar conceptual choice point. Chapter 6 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) and *climbing* (a cluster of ways of viewing the universe as just a part of another). I'll do my best to make these mathematically tricky ideas palatable to philosophers, and explain some of their philosophically relevant features. Chapter 7 will argue that the iterative conception of set is defective in certain ways, in particular it fails to answer all the questions we want. Instead, I argue, there is a kind of conception latent in much set-theoretic thinking that I call the *absoluteness conception*. The core idea at the heart of the absoluteness conception is that any set that 'could' exist does exist. Chapter 8 will explain a paradox for this conception, and explain how it is linked to the incompatibility of constitutive principles (much like we saw with the naive conception). Chapter 9 will suggest that we can view two conceptions of set at play in contemporary set-theoretic discourse (what I call the *forcing absoluteness* and *climbing absoluteness* conceptions) and argue that, much like our intellectual ancestors, we are at a conceptual crossroads *right now*. Chapter 10 will provide some objections, and importantly identify the limitations and somewhat simplistic nature of the study in this book. Chapter 11 will provide a short concluding summary, but more importantly will identify some further work that is needed in order to clarify the space of conceptual choices. I hope that the reader comes away with a sense of how set theory is philosophically interesting, and the vastness of the conceptual space offered by reflecting on our set-theoretic concepts and conceptions.

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 difficult mathematical ideas, *this is not a textbook*. My approach is one of conveying underlying ideas, rather than giving everything in full rigorous detail. Readers who wish to go through the details are encouraged to follow the citations to what I consider to be good expositions of the relevant formal details.

Second, the pacing of this book will feel slightly odd. The two aims I have set for myself are fundamentally in *tension*, 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 in the early chapters, whilst outstripping what can be expected of an early student (however talented) later. To combat this, 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 later material as accessible as possible. This said, some of the material in this book is *hard* if you aren't familiar with the relevant bits mathematical logic. My aim is to make the difficult material *accessible* and not, per impossibile, *easy*.

Third, 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). Throughout, we'll be talking about *concepts* and *conceptions* (see Ch. 3 for this distinction) and from now on I denote concepts by caps (e.g. FAIRNESS is the concept of fairness) and conceptions are enclosed within guillemets (so I'll talk about the «iterative conception» of SET). 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.

Finally, it's important to note that in parts I'll speak as though things are conceptually neat and tidy in various ways, and/or admit of clear answers. By and large I think the conceptual landscape is far messier than these pages indicate (we'll discuss this more in Chapter 10). However, just as it can be helpful to consider illustrative simplifications in many areas of science, the same goes here. It's a useful exercise to consider sharpenings of what is probably a messy landscape in order to see some of the contours better, and figure out what we want to say about the simplified case in order to illuminate the complicated one. With this in mind, let's get ready for our journey.

Chapter 2

Why set theory?

Before we start getting into the conceptual 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 and conceptions of set can have. These will be important later when we come to assess the value of different conceptions.

What are sets? The idea of what we should take our term "set" to denote, what our concept of SET is like, and the nature of the sets themselves will form the main focus of this book. So we should start with a rough and ready definition:

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?** Why has such a seemingly banal notion of collection commanded so much attention in philosophy? For all I've said above, sets just look like a convenient way of talking about a bunch of objects reified into an object-like collection, but its not clear why this deserves our attention any more than other kinds of object. So why should we care?

It's useful first to consider a bad answer (but one that helps us see the role of set theory more clearly):

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][1]

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 kinds of collection or 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[2], 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 can be taken *intensionally*, where identity is not taken to be governed by extensionality. 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, this wouldn't suffice as an *explanation* of why it has garnered so much philosophical attention. Presumably, collections of beer coasters aren't the most scintillating kind of object we can theorise about. So we need some stronger reasons to be so interested in set theory.

Here's what I take to be the core point about set theory: *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 has yielded two interlinked faces to set theory:

[1] 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.

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

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

Philosophical Example. Set theory provides a 'test case' for many problems of philosophy.

This division is far from exclusive. Certainly there are cases where we might think that set theory and philosophy are inextricably intertwined.[3] Indeed, this book can be viewed emphasising the act 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 of interest or is intertwined with philosophy, and set theory studies its own properly mathematical questions. However the distinction serves as a rough guide to the different facets of set theory that will be interesting to philosophers.

At this stage, we'll keep things relatively informal, but a little precision will be helpful. The particular set theory we'll be concerned with in this book will be *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, find the range of any function you want).

Recently, Penelope Maddy has isolated some *mathematical goals* of set-theoretic foundations built on ZFC in [Maddy, 2017] and [Maddy, 2019]. I'll provide some examination of Maddy's ideas, and I'll suggest some modifications and additions of my own.

Earlier I mentioned that set theory is a powerful device of *representation*, and many of the interesting goals and desiderata of set theory are linked to this idea. We can encode all mathematical objects by sets.[4] What do I mean by 'encode' 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 on Euclidean space, we can represent this straight line by some function $f(x) = bx + c$, and think of the straight line as composed of *ordered pairs* $\langle x, bx + c\rangle$. This can help us when, for example, trying to compute the lengths of line

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

[4] 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]) for details.

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.

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[5], rationals as pairs of natural numbers, reals as Dedekind-cuts of rationals,[6], ordered pairs as Kuratowski-ordered pairs[7], and functions/relations by sets of ordered pairs[8]. Of course there's lots of choices, and this is just an illustration of *one* way you might do things.[9]

Using similar tactics, any mathematical object we have come up with can be encoded by sets. (There are some controversial cases like categories, and I leave a remark about these to a footnote.[10]) This has some important consequences mathematically speaking. 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 name—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 (but more on this later). This Maddy argues, gives us:

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

---

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

[6]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.

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

[8]So the function f is encoded by $\{\langle x, y\rangle | f(x) = y\}$.

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

[10]There is an objection to set theory that 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.) I think the study of proper-class-sized objects is a perfectly good 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).

The thought here is that because we have **Generous Arena** and can view mathematical objects as encoded by sets, a proof about a mathematical object can be regarding as correct if it could in principle be translated into a proof in set theory about properties of the relevant mathematical code(s). Of course, the "*in principle*" is important here—outside of set-theoretic mathematics it is very clunky to work with these codes, and we shouldn't expect mathematicians to actually go about their day to day 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 this kind of foundation 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.)

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.

There are 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 no element of x has infinitely many R-predecessors. This notion is important, since it helps us think of performing *actions* or *operations* into the infinite along R, since there are only *finite* descending R-chains. Within ZFC one can represent and develop an arithmetic for these orders, defining notions of *ordinal addition*, *multiplication*, and *exponentiation*.[11] This provides us with ways of generalising normally finite operations (e.g. computation) into the infinite (e.g. *infinite* time Turing machines[12]).

Cardinal numbers, by contrast, can be thought of as answers to the question of how *many* objects there are in a particular set. In particular, by letting two sets X and Y *have the same cardinality* iff there is a bijection $f : X \rightarrowtail\!\!\!\!\rightarrow Y$ (and representing cardinals by particular kinds of sets) ZFC provides a theory in which the cardinal sizes of any sets can be compared, and natural operations like multiplication, addition, and exponentiation generalised and computed.[13] The success that ZFC

[11]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*.

[12]See [Hamkins and Lewis, 2000].

[13]Again, there's a variety of ways one might proceed, but here's a typical one. The

is striking—seemingly giving finite beings (e.g. us) the ability to reason about large infinite objects. Many surprising facts can thereby be shown. For example, we can prove that:

Theorem 2. There are as many natural numbers as there are squares of natural numbers.

This is, pre-theoretically, hugely surprising since the squares of n and $n + 1$ get more and more spread out as n gets larger. Indeed, the intuition behind this theorem was even regarded as a kind of 'paradox' by the likes of 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.

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. $\mathbb{R}^n$).

This ability to calculate the relative sizes of sets is an astounding achievement for mere finite beings like ourselves. It also helps us understand the role of infinity in certain arguments. Here, for example, is a debate that occurs in the context of the philosophy of economics:

Debt is unfair. Suppose that we loosely think of *debt* as the transfer of resources from future generations to present ones. Consider an infinite line of people, each holding a bottle of beer, with one person (the present person) at the front. Each person hands a bottle forward, until the person at the front has two bottles and everyone else has one. Thus, debt is unfair—we enrich ourselves by taking from future generations (even though they are not thereby made poorer from it).

NAB: Is this example needed/helpful/necessary?

Those familiar with set-theoretic mathematics, and in particular the example of Hilbert's Hotel, will notice a flaw in this argument—we could set things up so that everyone ends up with infinitely many

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.

beer bottles, if only we do so cleverly.[14] What this shows is not necessarily that debt *isn't* unfair, but that introducing infinity naively into certain examples can result in unintended or surprising consequences and should be treated with caution.[15,16]

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 (codes of) real numbers is greater in size than the cardinality of the (codes of) natural numbers, in the sense that there is no bijection between the domains of $\mathbb{N}$ and $\mathbb{R}$ (but clearly every natural number can be thought of as a real number).[17]

The problem goes far deeper. 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.[18]

Again, Cantor's Theorem is striking. It seems to imply, on the basis of natural principles about sets, that there are a *never ending hierarchy*

[14] Here's how: Have every other person give a bottle of beer (everyone has at least one) to person at the front of the queue. While this is going on, out of everyone that isn't giving to the person at the front of the queue, have every other person give a single bottle to person number two. Have every other person out of the people not giving to person number two, give a single bottle to person number three, and so on. After one round of giving, everyone has infinitely many bottles of beer.

[15] I am grateful to Alexander Douglas for discussion of this example, and directing me to some discussion of this on the following blog posts: PROVIDE CITATIONS.

[16] Another example, due to Barbara Montero and Joel David Hamkins, concerns how infinite assumptions can mess with utility. 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. For 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 sphere 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 converse disposition). However, Montero and Hamkins argue, we should *prefer* to be in the expanding happiness world (since then we just have to wait long enough to be blissfully happy forevermore). [NAB: Check the example.]

[17] I won't discuss a proof of this since it's available in many textbooks, and we'll discuss the more general version of Cantor's Theorem later.

[18] We'll discuss a proof of Cantor's Theorem later, in particular as it relates to the paradoxes in Chapter 4.

of infinite sets, since the power set of any set x is always bigger than x. How on earth can such seemingly innocent principles yield such apparent ontological profligacy?

We'll discuss some of these ideas throughout this book. For now let's note that Cantor's Theorem produces much of the interest of cardinal arithmetic—whilst addition and multiplication are trivial for 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.

This success, however, 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.

Before we discuss the details let's remark that the *mere fact* of independence 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. Let's start by considering the following hypothesis:

Definition 7. Let's have 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 αth cardinal number. A routine argument shows that $2^{\aleph_0} > \aleph_0$ (by Cantor's Theorem). But is there anything in between? That is, does $2^{\aleph_0} = \aleph_1$ (this is known as the *Continuum Hypothesis* or CH)? Or are there any cardinalities in between, and in fact $2^{\aleph_0} > \aleph_1$?

As it turns out, both CH and $\neg$CH are both consistent with ZFC (assuming ZFC itself is consistent). We'll explain how this works later (Chapter 6).

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 state within ZFC the sentence that there's no proof of $0 = 1$ derivable from the axioms of ZFC). Call this sentence Con(ZFC). But now we can point to:

Theorem 8. (*Gödel's Second Incompleteness Theorem*) Assuming that ZFC is consistent[19], then $Con(\mathsf{ZFC})$ is not provable within ZFC, and nor is $\neg Con(\mathsf{ZFC})$.

Within set theory we can study a wide variety of sentences that have different consistency strengths—one can prove other theories consistent from ZFC with the sentence added. As it turns out, CH and ¬CH are not like this (ZFC, ZFC+CH, and ZFC+¬CH are all *equiconsistent*). Obviously adding $Con(\mathsf{ZFC})$ results in a consistency strength increase. There is another kind of axiom—so called *large cardinal axioms*—that are important here, serving as the natural indices for consistency strength. One kind of large cardinal postulates the existence of sets with a lot of *closure* properties. Here's an example:

Definition 9. A cardinal κ is *strongly 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 $f : x \to \kappa$, the range of f is bounded by some $\gamma < \kappa$ (here, we say that κ is regular).

That's a lot of formal details, but it's instructive to think about what such an axiom says. Such a κ seems very big—clause (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. Interestingly, you can show that an inaccessible cardinal κ suffices to produce a model for ZFC, and so by Gödel's Second Incompleteness Theorem (and the Completeness Theorem) you can't produce an inaccessible cardinal from ZFC alone. Set theory has discovered a whole hierarchy of these cardinals with greater and greater closure properties (one can, for example, insist that κ have κ-many inaccessibles under it).

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. There are more kinds of independence (for example where we consider strong axioms that don't directly postulate the existence of cardinal numbers with

[19]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.

closure) but we'll not consider them just yet since we've got enough to chew on.

We should pause for a moment to reflect on what this independence tells us about the nature and limits of human thought, at least insofar as what we get from ZFC. Whilst ZFC does give us the resources to compute the cardinals and 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. We can therefore isolate the following philosophical aspect of set theory.

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.

When all is said and done, however, set theory constitutes our main theory 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). This yields another foundational goal of set theory:

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

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 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.

At this point, one might wonder: "Why should I be so concerned with consistency here?" 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 views, such as concerning mereology (e.g. [Uzquiano, 2006]). However, I also want to point out 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 and this constitutes a kind of 'paradox' (perhaps in a weak sense). Some of these we'll see later, and some others I mention in a footnote.[20] 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, and the fact that set theory gives us the tools to locate and diagnose these inconsistencies helps us elucidate our **Limits of Thought** and also gives 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 Examples**, 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 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

[20] For example, the embedding template of embeddings $j : V \to M$ 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 (cf. [Reinhardt, 1974], [Koellner, 2009]), and modal reflection principles are more flammable still (see [Roberts, 2019]).

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 out **Theory of Infinity**). Thinking about the above in this dual light will help to illuminate some of the issues and arguments that will crop up later.

Chapter 3

Concepts, conceptions, and conceptual engineering

I hope the reader now has a picture of why set theory is important for the philosopher and what some desiderata on a concept of set and set theory might be. I hope to make the importance of some of the points outlined in the previous chapter clearer as we move on.

This book will explain how *concepts* and *conceptions* of set can change, and what constrains their change. For this reason, some of the literature on *conceptual engineering* becomes especially relevant. This chapter will examine some of the views out there and set the stage for the rest of the book. In particular, I want to set up the notions of *concept*, *conception*, and the field of *conceptual engineering*. Let's deal with each of these in turn.

3.1 Concepts

Let's start with *concepts*. We should start with a caveat: What exactly we mean by the term "concept" and how we should understand their metaphysics is an old and difficult question. There are multiple answers we might give (e.g. they could be understood as mental representations, dispositions/abilities, or abstract objects). I certainly don't have enough space to address these difficult questions here. Nor do I think that philosophers and cognitive scientists yet have a satisfactory answer to these questions. The difficulties surrounding many classical theories of concepts are visible in the vast body of literature, but an interesting (although tricky) presentation of some of these issues is explored by Mark Wilson in [Wilson, 2006].

There are, however, some things we *can* say about concepts that I think suffice for the task at hand (namely analysing set-theoretic concepts/conceptions), even if the interesting metaphysical question has to be left unanswered. Concepts are the kinds of thing that figure in our representation of the world. It may even be that "concept" isn't the best term to use here (this is suggested by Cappelen [provide citation]). Nothing turns on this for us, I just mean our devices of mental representation. Important for us will be that they are the sorts of things that have *constitutive principles* where a rule or condition is *constitutive* for a concept when it (partly) determines the meaning of the concept and conceptual identity (assuming that there is such a thing).[1] So, for example, I can talk about the concept DOG as having the following constitutive principles:

- If x is a DOG, then it is a member of the biological species *canis familiaris*.
- Paradigmatic dogs have two eyes,
- are quadrupedal,
- have fur,
- and so on...

Note that the constitutive principles for DOG don't have to *all* be satisfied (there are three-legged and one-eyed dogs for example), but if a sufficient number are violated then we would not regard the relevant object as a dog.

Constitutive principles are important as a lack of agreement on constitutive principles between speakers can work as an interpretive 'red flag' that speakers do not mean the same thing by the use of their words.[2] If for example, you point to Story the cat and say "What a lovely dog!" I am likely to regard you as either very confused or using the word "dog" in a different (non-standard) way. So constitutive principles for concepts are how we know that we have a degree of semantic agreement, and can go on to discuss properties/constitutive principles that the relevant concept might have.

3.2 Conceptions

Closely related to concepts are *conceptions*. This is a notion that has been in the air in philosophy for a while. For instance, we've already

[1] Here I am mostly following [Scharp, 2013] and [Incurvati, 2020].

[2] See here [Scharp, 2013], p. 50.

mentioned the «iterative conception» of SET a few times. In many ways, this book will focus on the nature of concepts and conceptions, and how they change. (I'll include both when I talk about "conceptual change".) But what is a *conception* and how is it different from the notion of *concept*?

Recently [Incurvati, 2020] has provided a useful account of conceptions, and it's the one we will take up here. A conception is a particular way of sharpening what we take to fall under a given concept. Incurvati puts the distinction as follows:

Conception. A conception of C, where C is a concept, is a (possibly partial) answer to the question "What is it to be something falling under C?" which someone could agree or disagree with without being reasonably deemed not to possess C.[3]

Here's an example provided by Incurvati.[4] Suppose someone is going to be rewarded over someone else by their company for their work on a case (let's say there was a good outcome but the person did not put much work in). Jane and Susan disagree over whether this decision is fair, Susan thinks companies should reward employees on the basis of outcomes, whereas as Jane thinks that companies should reward employees on the basis of effort.

In this case, we may assume that both Jane and Susan possess the concept of FAIRNESS. But they have different *conceptions* of fairness—they disagree on the extension of the concept FAIRNESS. In particular Susan has what we might call the «fairness-by-effort conception» of FAIRNESS, whereas Jane has the «fairness-by-outcome conception» of FAIRNESS. A *conception* can thus be thought of as *adding* to the constitutive principles of a concept in such a way that the violation of those constitutive principles does *not* constitute an 'interpretational red flag' for the concept in question.

One issue (not *directly* considered by [Incurvati, 2020]) is that the notion of what is a concept or a conception admits of a relative reading.[5] For instance I can talk about the «fairness-by-effort conception» of FAIRNESS, or the «fairness-by-outcome» conception of FAIRNESS. But I can *also* talk about the *concept* of FAIRNESS-BY-OUTCOME, where

[3]See [Incurvati, 2020], p. 13.

[4]See Ch. 1 of [Incurvati, 2020].

[5]Whilst [Incurvati, 2020] doesn't *directly* consider this relativisation, I do think something like it is implicit in some of the things he says, for example when discussing the addition of maximality ideas to the «iterative conception» of SET (see esp. Ch. 8). The idea of relativising the concepts/conceptions distinction emerged in discussion with Øystein Linnebo, and I'm grateful for his input into the ideas here.

someone can succeed or fail in understanding the constitutive principles of what FAIRNESS-BY-OUTCOME entails. By doing so, I fix the constitutive principle that rewards should be determined by outcomes. We can then, in turn, discuss *conceptions* of FAIRNESS-BY-OUTCOME. Here's an example:

> Let's suppose that Mar is rewarded for making their company a good deal of money, but had to do something socially problematic in the course of doing so. Anwar and Bo disagree on whether this is fair. Both Anwar and Bo use the concept of FAIRNESS-BY-OUTCOME, but Bo has the conception that FAIRNESS-BY-OUTCOME should be understood in terms of making the company money, whereas Anwar thinks it is the benefit to society as a whole that is important.

In this case, I contend, Anwar and Bo have different conceptions of FAIRNESS-BY-OUTCOME. We might say that Anwar has the «societal-benefit conception» of FAIRNESS-BY-OUTCOME whereas Bo has the «revenue conception» of FAIRNESS-BY-OUTCOME. And we might go even further, considering conceptions of the concept FAIRNESS-BY-REVENUE-OUTCOME and so on.

What we are doing with this talk is demarcating what constitutive principles we are taking to be fixed for a particular device of mental representation (i.e. a concept or conception). We can have many different conceptions of a concept, but the constitutive principles that are added by a conception can be taken to be fixed when moving to the concept correlated with that conception.

This analysis of concepts and conceptions has some useful features. First, it accounts for how constitutive principles that begin life as part of a conception of a concept can slowly form part of the concept attaching to our use of the term. A nice example here comes from [Shapiro, 2013]. It's highly plausible that our concept EFFECTIVELY COMPUTABLE used to merely contain the rough-and-ready constitutive principle that there should be a procedure that could be followed for computing an answer to the relevant problem. One conception of EFFECTIVELY COMPUTABLE that emerged was the «Turing computable conception» of EFFECTIVELY COMPUTABLE, as given in [Turing, 1937]. It's highly plausible that this was initially a non-trivial *conception* of EFFECTIVELY COMPUTABLE. In light of the many computational equivalences that were subsequently observed and the resulting Church-Turing Thesis, however, for many theorists, the concepts EFFECTIVELY COMPUTABLE and TURING COMPUTABLE just have the same constitutive principles—many people just mean Turing computable when they use the term "effectively computable". So what was originally a conception can become the default concept attaching to a term through time.

Second, it allows those who have a differing conception of a concept to engage with those who differ. In our example of FAIRNESS, Jane differs from Anwar and Bo in that she thinks that the «fairness-by-effort conception» of FAIRNESS is correct. However, she can engage with Anwar and Bo's disagreement, since presumably she possesses the *concept* FAIRNESS-BY-OUTCOME (it's just that she thinks that the «fairness-by-outcome conception» of FAIRNESS is wrong). Indeed, as honest philosophers, we often *have* to take some constitutive principles we do not agree with as fixed when discussing a concept on its own terms.

Let's briefly remark how this might work in the case of *set theory*, and identify the route we'll be taking in the rest of the book. Obviously we're just working with a template right now, since the details are yet to come! But let's start with the concept of COLLECTION. As remarked in Chapter 2, there's lots of different ways we can think of collections (they might be extensional or intensional, objectual or non-objectual etc.). But as we saw, one conception of COLLECTION is the «set-theoretic conception» of COLLECTION—collections are viewed as extensional entities that are objects over and above their elements. Now we have a particular conception of COLLECTION—the «set-theoretic conception» —and we can talk about SET as a concept in its own right. It's natural then to talk about *conceptions* of SET.[6] As we'll see, we have the «naive conception» , the «stratified conception» , and the «iterative conception» (as well as more besides!). And, as we'll see, there's then no obstacle to considering the *concept* ITERATIVE SET, and what conceptions of *this* might be like, and what happens when we take those conceptions as concepts in turn.

3.3 Inconsistent concepts and conceptions

One feature of this way of looking at concepts and conceptions is that they can be inconsistent. The following example is due to Kevin Scharp:

Definition 10. x is a *rable* is given by the following application conditions:

(i) RABLE applies to x if x is a table.

[6] Indeed this is the strategy of [Incurvati, 2020], which we'll discuss a lot.

(ii) RABLE disapplies (i.e. does not apply) to x if x is a red thing.[7]

This is an inconsistent concept at the present world. Plenty of things fall under it—for example the white table I am writing at passes—however one encounters a problem when confronted with a red table.

The example of RABLE is gerrymandered, but there are several interesting scientifically useful concepts that one might argue are inconsistent. Some that have been proposed include (Newtonian) MASS, the early analysts' concept of DERIVATIVE, and TRUTH.[8] Despite the historical significance of these concepts, they have been argued to be inconsistent (and hence defective). How should we proceed given this predicament?

3.4 Conceptual engineering

The natural answer is to *change* our concepts. The fact that I've talked about moving between concepts and conceptions, and that conceptions can be taken as concepts in their own right is highly suggestive of the idea that I think that our concepts and conceptions can change, and that what we do with out word "set" might also change. Indeed, this will be a cornerstone of this book. In this regard it's useful to say a few words about *conceptual engineering*, which we can roughly characterise as follows:

Conceptual engineering is (roughly speaking) the field that concerns itself with the (1) design, (2) implementation, and (3) evaluation of our concepts and conceptions.[9]

Before we continue, we should make some remarks about the place of conceptual engineering in philosophy and the role of (1)–(3).

First, conceptual engineering has become somewhat fashionable lately, but has been around for a *long* time (this is a fact to which many scholars of conceptual engineering are sensitive). In fact, philosophers have pretty much always asked questions about our concepts, what their constitutive principles might be, and how they can be evaluated. One can, for example, read much of Plato in this way if one so desires (e.g. the *Meno*, *Theaetetus*, and *Euthyphro* all spring to mind, and examples can be multiplied). Often Plato's dialogues involve Socrates

[7] I will use "disapplies" following [Scharp, 2013] as the antonym for "applies". For this example see [Scharp, 2013], p. 36.

[8] Some of these are considered by [Scharp, 2013].

[9] This definition is adapted from [Chalmers, 2020].

requesting that an interlocutor provide constitutive principles for the concept in question, before evaluating their answer (almost always in the negative). It is, however, with the work of Carnap, and in particular his work on *explication*, that conceptual engineering is often thought to really take off.[10] For Carnap, explication involved *sharpening* concepts—in our terms, adding constitutive principles to form a preliminary conception before adopting those principles into the concept.

For many of the ideas we'll be considering here, Carnapian explication suffices. However it should be noted that it's not the only kind of conceptual engineering we might engage in. To see this, let's discuss (1) to (3). (These are, loosely speaking, taken from [Chalmers, 2020].) It should be noted that the ordering on (1) to (3) is somewhat arbitrary, they can occur in any order you like and often concurrently, but speaking of an ordering helps to get the rough idea going. In stage 1 we design a concept. This might be through proposing a conception of a concept (as in the case of our example of FAIRNESS) or completely *de novo* (plausibly the example of SUPERVENIENCE is one from philosophy, and STRING from theoretical physics). Call this the *design stage* of conceptual engineering. In stage 2, we encourage the uptake of our particular concept or conception, and for many cases we can view this as the time in which a conception becomes a mainstream and used concept. Call this the *implementation stage*. As many philosophers have noted, the implementation stage may require significant *activism* in order for the concept/conception to be taken up (CITE CAPPELEN). Stage 3 we call the *evaluation stage*, here we take concepts/conceptions (whether implemented or not) and assess whether they are fit for purpose, and in what respects they are defective (inconsistency being a good example of defectiveness).[11]

If a concept or conception is judged defective upon evaluation, several options are open to us. One is to propose a re-engineering of the concept, and enter the design stage again. With concepts that are vague in some way, it's natural to perform Carnapian-style explication, proposing conceptions of the concept that add constitutive principles. But this need not be the case every time we re-engineer. We could also propose a *coarsening* of the concept in question. This has occurred, for example, with the concept MARRIAGE—a concept that at

[10]For the history of explication as it appears in Carnap's work, see supplement 'D. Methodology' to the Stanford Encyclopedia article on Carnap [Leitgeb and Carus, 2021].

[11][Scharp, 2020] even goes so far as to argue that philosophy is *predominantly* the study of inconsistent concepts.

one stage might have contained the constitutive principle that it had to be between a man and a women, but which has been significantly ameliorated by *removing* this constitutive principle and coarsening the concept. Another possibility for engineering after evaluation is to *abandon* the concept (and use of the term) altogether, as has largely happened within the scientific community regarding the concept PHLOGISTON.

Along with conceptual engineering we find the question of how lexical terms like "set" are meant to match up with concepts like SET.[12] I don't, for example, prove the Twin Prime Conjecture by having PRIME be the conceptual correlate of the lexical item "pair of twin primes" (there being, of course, infinitely many prime numbers). Whether two speakers mean the same by a lexical term, as we've suggested, will depend on the level of agreement between their constitutive principles, and whether they should thereby be taken to be disagreeing on conceptions of a concept or simply meaning something different by the terms themselves.

So, that's the idea of a distinction between concepts and conceptions, and how we might integrate these ideas with the wider literature on conceptual engineering. Of course, there is a huge amount that one could dispute here, and the fine details are tricky to make out. Nonetheless, this coarse outline is enough to make some progress on the case of set theory, and in particular see (1.) How conceptual engineering has happened in set theory, and (2.) How it might go in the future.

[12] See, for example, [Chalmers, 2020], p. 8 of preprint on website, [FIND CORRECT CITATION].

Chapter 4

The naive conception of set and the classic paradoxes

We've now got some desiderata for set theory on the table (Chapter 2) and an idea of how concepts/conceptions can be engineered (Chapter 3). 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 want to give a 'new' presentation of the classic paradoxes that has been proposed recently by philosophers and examine a particular diagnosis that will help to make some of the more difficult material towards the end of the book easier to follow.

4.1 What do we want from a conception of set?

One way of approaching the issue of the set-theoretic paradoxes in light of our focus on conceptual engineering is through the following question:

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 an understanding of formal axiomatic theories, knows what a many-sorted and one-sorted theory is, and knows the difference between first- and higher-order logic, and has at least a passing familiarity with the model theory for these languages. When and where an unfamiliar notion pops up, the reader can consult

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

[Button and Walsh, 2018] for an excellent philosophical examination of philosophy, axiomatic theories, and their model theory.

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 a 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 (i.e. plausible constitutive principles) 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 the «iterative conception», stage theory, and ZFC), but equally they can also be something more informal.

4.2 The naive conception of set

Our first such conception will be the «naive conception» of SET:

Definition 11. (Informal) The *«naive conception»* of SET adds the constitutive principle 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]

The «naive conception», as a conception of SET, clearly motivates adoption of the *extensionality axiom* (which says that any two sets with the same members are equal), but also motivates:

Definition 12. The *Naive Comprehension Schema* asserts that for every one place formula $\phi(x)$ in the language of set theory (i.e. the language that has $\in$ as its sole non-logical symbol), 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 with Extensionality. Let's see how.

4.3 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

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

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

why we'll look at them:

(1.) Part of what we will see later is a 'new' kind of 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 have been some developments 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*.

In this book, I'll only really consider Russell's Paradox and Cantor's Paradox. The Burali-Forti Paradox is interesting, however it is complicated by the fact that one has to use set-theoretic codes for the ordinals themselves (conceived of as order types) in order to make the paradox bite.[4] With this caveat in mind, here's simple presentations of the paradoxes.

Russell's Paradox. Consider the condition $x \notin x$. By Naive Comprehension, this determines a set y. But then $y \in y \leftrightarrow y \notin y$.

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. Clearly there is a bijection (i.e. a one-to-one and onto function) $f : u \rightarrowtail\!\!\!\!\rightarrow \mathcal{P}(u)$, since every member of $\mathcal{P}(u)$ is also a member of u. Now consider the set $r = \{s|s \notin f(s)\}$. Again, we have $r \in r \leftrightarrow r \notin r$. 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 bijection $f : x \rightarrowtail\!\!\!\!\rightarrow \mathcal{P}(x)$.

So far, so well-known. Many introductory textbooks contain a presentation of the paradoxes. However, something philosophers have realised recently (and something that will prove to be important later) is that each of these paradoxes can be viewed as theorems about the (non-)existence of certain kinds of *function*.[5] In particular we give the following presentation:

[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, [Meadows, 2015], [Incurvati, 2020], [Scambler, 2021], and [Builes and Wilson, 2022].

The Cantor-Russell Paradox. Define u and $\mathcal{P}(u)$ as in Cantor's Paradox. Clearly there is an *injection* $f : \mathcal{P}(u) \rightarrowtail u$ since the identity map $f(x) = x$ will do. Now we consider the set $r =_{df} \{y|y \notin f(y)\}$. And again we note that $r \in r \leftrightarrow r \notin r$.

The important thing to note is that in this context (where f is the identity map) the contradictory set r we get is *both* the operative problematic set for Cantor's Theorem/Paradox *and* the ordinary Russell set (since f is the identity map here, the set $\{y|y \notin f(y)\}$ *just is* $\{y|y \notin y\}$). So Russell's Paradox and Cantor's Theorem/Paradox are not just *superficially* similar, but in many contexts come down to definition of *exactly the same set*, and the core issue is the (non-)existence of an injection $f : \mathcal{P}(x) \rightarrowtail x$.

This observation works in the other direction to, where we assume that we have an *onto* function $f : u \twoheadrightarrow \mathcal{P}(u)$. Without loss of generality, again this can be the identity map (since if $x \in u$, every member of x is in u, and so $x \in \mathcal{P}(u)$). Now we can just consider the set $\{y|y \notin f(y)\}$ as before, and so the existence of a *surjection* gets us both the set for Cantor's Paradox and the Russell set.

4.4 Diagnosis

So, as we all know, 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] and [Incurvati, 2020]. We'll follow Incurvati's presentation here, since it will be instructive for making comparisons with conceptual engineering (and, in particular, [Scharp, 2013]'s examination of TRUTH).

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

Universality. A concept [or conception] C is universal iff there exists a set of all the things falling under C. (adapted from [Incurvati, 2020], p. 27)

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:

Indefinite extensibility. A concept [or 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. ([Incurvati, 2020], p. 27)

Indefinite extensibility also follows from the Naive Comprehension Schema, since any time we have a set x the Naive Comprehension Schema gives us the juice to use the Cantor/Russell reasoning and diagonalise to find a set not in x (e.g. one of the members of $\mathcal{P}(x)$).

Clearly, any concept/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 concept/conception. So, what might our response to this state of affairs be?

Chapter 5

Early engineering

This is the situation we found ourselves in at the turn of the 19th century. The burgeoning field of set theory was clearly *useful*, but the «naive conception» was *deeply* flawed. In this chapter, I want to present the emergence of the familiar «iterative conception» of SET, as well as the somewhat *less* familiar «stratified conception» . Both of these are considered by [Incurvati, 2020], but I also want to link these to the literature on conceptual engineering. Analysis of some of the history here in terms of concepts, conceptions, and the shift from a conception to the incorporation of a dominant concept, helps to make sense of the intellectual landscape various agents have inhabited. In particular, I'll argue that there was a period where we moved from the «naive conception» to the «iterative conception» and «stratified conception» (but where it was not clear which would become dominant), before widespread *acceptance* of the «iterative conception» as the better of the two. The end result, I argue is that now the concept correlated with the use of many mathematicians' use of the lexical term "set" is now ITERATIVE SET.

5.1 The stratified conception of set

What we need is a conception of set that (1.) provides an account of what sets are like, (2.) motivates as nice a possible theory (keeping in mind the constraints of Chapter 2), and (3.) is consistent. Let's meet our first contender:

Definition 13. (Informal) The «*stratified conception*» of SET holds that sets are the extensions of *well-defined* predicates, where a predicate is *well-defined* iff it respects typing restrictions.

Let's say a little more about what is meant by *typing restrictions*

([Incurvati, 2020] provides a good examination of the history, but we'll just take the version provided by [Quine, 1937]). We start with the following definition:

Definition 14. 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.

We then can formulate:

Definition 15. The *Stratified Comprehension Schema* asserts that for all **stratified** formulas $\phi(x)$, there is a set of ϕs, i.e.

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

The theory that adds every instance of Stratified Comprehension to Extensionality was developed by [Quine, 1937] and is known as *New Foundations* (NF for short). There is no widely accepted proof of the consistency of NF, though there are a couple of attempts whose status as proofs is unknown. The addition of urelements, and subsequent weakening of Extensionality to imply that only *sets* with the same members are identical, yields a system known as NFU that is consistent relative to Peano Arithmetic.

Putting aside these details, we might ask whether or not there is a cohesive idea behind the «stratified conception» . Much can be said about this and there's lots of nuance that we'll put aside (see [Incurvati, 2020] for some further details). However, it's important to see that there is a legitimate *conception* in play here, and not merely some ad hoc restriction.

One idea is to hold that the «stratified conception» is based on the notion of *objectified property*. This approach is suggested by [Incurvati, 2020], drawing on work by [Cocchiarella, 1985]. The idea here is that sets are what you get when properties are associated with objects. On this conception, some properties can have first-order surrogates, and some can't. And all the properties resulting from stratified formulas can have surrogates. A similar proposal is made by [Holmes, 1998] who thinks of sets as 'names' for aggregates of names, with the Russell reasoning showing that not every aggregate can be named, and good names needing to be stratified. One way of making this precise is to think of sets as data types and stratification stemming from a need for implementation independence.[1] Another idea is to regard sets as kinds of facts, and the need for stratification as coming from the Vicious Circle Principle.[2] All of these attempts to motivate Stratified Comprehension view sets as somehow first-order correlates of higher-order entities. In particular, they can be thought of what one gets by collapsing the type-theoretic hierarchy to sets. So, even if there are questions about the «stratified conception» as a bona fide conception, there's at least some reasonable stories that we can provide on its behalf.

These stories are somewhat complex, and the particulars are quite tricky. Thankfully these details aren't central to what I'll argue here, and we needn't get bogged down in them. What I rather want to do is draw the reader's attention to how the proponent of the «stratified conception» blocks the paradoxes. We noted earlier the difference between **Universality** and **Indefinite Extensibility**. The proponent of the «stratified conception» comes down on the side of rejecting **Indefinite Extensibility** and accepting **Universality**. The latter clearly holds, since the formula $x = x$ is stratified, and hence there is a universal set. However, **Indefinite Extensibility** fails. In particular the universal set is a set that cannot be extended and the standard reasoning for Cantor's Theorem fails in NF.[3] So when faced with inconsistent constitutive principles we can move forward by rejecting one of them (so long as we can provide a reasonable underlying conception). This idea is one that we will return to throughout the present study.

5.2 The iterative conception of set

As the reader may be able to guess, we'll now consider the «iterative conception» of SET, which takes the other available route of accepting **Indefinite Extensibility** but rejecting **Universality**.

We'll keep the idea of the «iterative conception» rough and imprecise to begin with (we'll make it sharper in just a tick, and we'll see why we're doing so later):

Definition 16. (Informal) The *«iterative conception»* of SET holds that 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 what are found at the stages.

[1]See especially Ch. 8 of [Holmes, 1998] and Ch. 6 of [Incurvati, 2020].
[2]For this approach, see [Hossack, 2014].
[3]For details see Ch. 6 of [Incurvati, 2020] and the references contained therein.

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. For our purposes this can be the empty set—nothing will hang on what we have at the initial stage. We then begin forming sets out of what we have using some given operations, and in this way obtain the sets.

The «iterative conception» of SET can in fact be split into two conceptions, a strong one and a weak one:[4]

Strong Iterative Conception of Set. 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.

Weak Iterative Conception of Set. The «weak iterative conception» of SET also holds that sets are formed in stages. Sets are formed by collecting together sets at previous stages. However we leave it open whether new subsets can be formed at subsequent stages. There are no other sets beyond those obtained this way.

Clearly any structure satisfying the «strong iterative conception» also satisfies the «weak iterative conception» . Let's see an example of the difference by going into more detail on each.

The «strong iterative conception» is very much the standard account provided. The idea that we get the sets by forming all possible subsets is akin to iterating the powerset operation. Unions of previous stages can be taken at limits.

This story can be made formally precise by adopting theory with variables for *sets* and *stages*, and primitives for *membership* (that holds between sets), *preceding* (that holds between stages), and *being found at* (that holds between sets and stages). Such an axiomatisation is given by [Button, 2021], drawing on work of Scott, Potter, Montague, Derrick, Boolos, and Shoenfield. Button shows that this stage theory is very closely related to a weak theory of sets that he calls *Level Theory* (they prove all the same sentences of *set* theory). This situation extends to ZFC, where natural theories of stages prove all the same set-theoretic sentences as the stage theory.[5] Moreover, within ZFC we can find correlates of the stages. The time has come to be precise about ZFC and what can be done there.

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

[5]For an easy presentation of these results, originally due to [Shoenfield, 1967] and [Boolos, 1971], see [Button, 2021], Theorem ????.

Definition 17. *Zermelo-Fraenkel Set Theory with the Axiom of Choice* (ZFC) 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) *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.

(v) *Axiom of Infinity.* There is a non-empty set such that if it contains a set z, it also contains z unioned with its singleton. The axiom thus guarantees the existence of an infinite set.

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

(vii) *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.

(viii) *Axiom Scheme of Replacement.* If a formula ϕ defines a function, then the image of any particular set under ϕ is also a set.

Using ordinal numbers, we can then (within ZFC) define:

Definition 18. *The Cumulative Hierarchy of Sets* is defined within ZFC as follows:[6]

(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)

We can then prove:

Theorem 19. (ZFC) For every set x there is an ordinal α such that $x \in V_\alpha$.

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

The core point is the following: The V_α *themselves* look like stages, and can be produced within ZFC. They are obtained by taking the empty set at V_0 (this is guaranteed by Infinity and Separation using the condition $x \neq x$), and iterating Powerset, collecting unions at limits (this latter operation is guaranteed by Replacement). So stage theory and 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. 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 stage theory is a *mathematical fact of life*—if you have ZFC you also have the «iterative conception» .

The «weak iterative conception» is in some ways less well studied, possibly partly because the strong account is seen as the default. It will, however, be important later. Recall that the core point about the «weak iterative conception» is that there is *no* guarantee that we get *all possible* sets formed at successor stages, as there is with the «strong iterative conception» .

Here's an example from set theory to illustrate the point. Often set theorists will talk about the *constructible universe* (or L) and *constructible hierarchy*. L is much like V, except that it is formed by taking *definable* powersets. A subset x of a structure $\mathfrak{M}$ is *definable* over a structure $\mathfrak{M}$ iff x is the unique object containing all and only the sets satisfying $\phi(y)$ for some condition ϕ in $\mathfrak{M}$.[7] Let's call the collection of all such definable subsets over a structure $\mathfrak{M}$ $Def(\mathfrak{M})$. Then L can be defined as:

Definition 20. 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.

Now the constructible hierarchy clearly satisfies the «weak iterative 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 coding new real numbers get formed as we climb through the first few stages above L_ω. So the L_α hierarchy does not satisfy the «strong iterative conception» , sometimes

[7]This is fiddly to formulate. Details are available in [Drake, 1974] [PAGE], [Kanamori, 2009] [PAGE], and [Jech, 2002] [PAGE]

there are possible subsets that don't get picked up when we move to a successor stage. Does this mean that the L_α hierarchy isn't properly *iterative*? I argue no, it is still iterative—there is still a perfectly good operation of set formation in play—it is just that it is not *strongly* iterative.

However: It should be noted that, assuming ZFC in the background, L also models ZFC, and so can build its own version of the V_α hierarchy, it's just that it's not the case that $L_\alpha = V_\alpha$ for every α, even if $V = L$. This distinction between weak and strong versions of the «iterative conception» will be important when we discuss the situation with respect to our contemporary conceptual situation later.

For now, let's identify that for *both* the weak and strong versions of the «iterative conception», **Universality** is out, but **Indefinite Extensibility** is validated. For any sets found at some stage, the set of all of them is (by the Russell reasoning) found at a later stage. So there is no set of all sets. So, the «iterative conception» takes the other conceptual road as compared to the «stratified conception».

5.3 The analogy with truth

Let's now contrast the above situation with what maybe a more familiar example for many philosophers from the conceptual engineering literature on truth. [Scharp, 2013] has argued that the concept TRUTH is inconsistent. The following principles, Scharp argues, are constitutive of TRUTH (letting Tr be the truth predicate):

Truth-theoretic ascent. $\phi \rightarrow Tr(\ulcorner \phi \urcorner)$

Truth-theoretic descent. $Tr(\ulcorner \phi \urcorner) \rightarrow \phi$.

As is well known, accepting both truth-theoretic ascent and descent will, along with classical logic and the material conditional, give you a contradiction very quickly. There is an enormous literature on solutions to the liar, but the solution proposed by [Scharp, 2013] is the following: We should replace truth with two related concepts ASCENDING TRUTH and DESCENDING TRUTH, where truth-theoretic ascent is constitutive for ASCENDING TRUTH but not DESCENDING TRUTH, and vice versa. This response, Scharp argues, allows ASCENDING TRUTH and DESCENDING TRUTH working in concert to fulfil the roles that we required of TRUTH, whilst having the substantial advantage of being consistent concepts. Much ink (or perhaps toner) has been spilled on this matter and I don't have time to address all the options here.[8] What I do want to do is point out the analogy here. In the case of NAIVE SET (or the «naive conception» of SET, if you prefer) we have two principles, **Universality** and **Indefinite Extensibility**, that collaborate to bring us to ruin. The «stratified conception» and the «iterative conception» each deny one of these two principles. Here I am simplifying issues slightly, since there's other options (e.g. the «graph conception» of SET). But this coarse description suffices to get the ball rolling for the similar moves we'll see later—when the constitutive principles of a concept or conception contradict one another, one can often move to a plurality of coherent concepts/conceptions by coming up with a story as to which should be dropped/accepted under the relevant concept/conception. The key point is that the new concepts/conceptions should be able to do the jobs (possibly separately) that we wanted the original inconsistent concept/conception for.

5.4 Engineering and the two kinds of set-conception

So, we've now seen that there are at least two kinds of conception of set that can come out of denying one of **Universality** and **Indefinite Extensibility**. I'll now present and add to some recent work arguing that the «iterative conception» (really the «strong iterative conception») became dominant, but we'll also discuss the place of the «stratified conception» and the intellectual climate philosophers and set theorists found themselves in on this journey. This will help us later to see our own intellectual predicament a little better.

Point 1. Early on, it wasn't clear what conception we had of SET. Sometimes reading the literature on the philosophy of set theory, one can get the impression that the «iterative conception» was latent in set theory even before it was fully isolated and interrelated with stage theory in the 1960s. However, this is decidedly not the case. As Michael Potter puts it:

> ...in an attempt to make the history of the subject read more like an inevitable convergence on the one true religion, some authors have tried to find evidence of the iterative conception quite far back in the history of the subject. [Potter, 2004, p. 36][9]

[8] In fact a special issue of *Inquiry* was released on [Scharp, 2013], see CITE INTRODUCTION TO VOLUME for a summary.

[9] Here Potter is contrasting the «iterative conception» with the «limitation of size conception » which we'll set aside for current purposes.

A history of the development of set theory is available in [Potter, 2004] and [Incurvati, 2020]. I want, however, to draw attention to the fact that several scholars did not have a precise distinction between the «stratified conception» and «iterative conception» , and their thought seems consonant with *both* at various points.

An oft-quoted passage of Cantor remarks that a set is a:

> ...many, which can be thought of as one, i.e., a totality of definite elements that can be combined into a whole by a law. [Cantor, 1883, p. 916]

This talk of a 'many' being thought of as 'one' through a 'law' is clearly more suggestive of something like the «stratified conception» , where we think of sets as given by conditions. On the other hand, the widespread acceptance of the Axiom of Choice by the likes of Cantor and Zermelo clearly meshes better with the «iterative conception» .

NAB: Maybe expand this last point? Is it strong enough, or should I find independent and explicit quotations from the likes of Cantor and Zermelo?

The problem extends much later in time. In 1917, Dimitry Mirimanoff wrote a paper entitled 'Les antinomies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles'. As well as considering Russell's paradox, he identifies what he sees as two kinds of sets:

> I will say that a set is *ordinary* just in case it gives rise to finite descents, I will say that it is *extraordinary* when among its descents are some that are infinite. [Mirimanoff, 1917, p. 42, my translation][10]

In modern terminology, Mirimanoff was distinguishing between well-founded sets (those who have only finite descending membership chains) and non-well-founded sets (those that have infinite descending membership chains). But if we want to say that the (strong) «iterative conception» is somehow latent in Mirimanoff's thinking, then Mirimanoff looks to be being rather silly here—if sets are found in a

[10]The original French reads:

> Je dirai qu'un ensemble est *ordinaire* lorsqu'il ne donne lieu qu'a des descentes finies; je dirai qu'il est *extraordinaire* lorsque parmi ses descentes il y en a qui sont infinies. ([Mirimanoff, 1917], p. 42)

well-founded hierarchy of stages by taking powersets, then of course the sets are well-founded.[11] (The issue of the «weak iterative conception» is somewhat more subtle, and we'll discuss this in Chapter 9.)

We can push this even later. For instance, the fledging of a full «iterative conception» is sometimes pinpointed to be [Zermelo, 1930]. But even then, as [Incurvati, 2020, p. 38] notes, Zermelo seems to suggest that there *could* be non-well-founded sets. Speaking about the Axiom of Foundation, he writes:

> This last axiom, which excludes all 'circular' sets, all 'self-membered' sets in particular, and all 'rootless' sets in general, has always been satisfied in all practical applications of set theory, and, hence, does not result in an essential restriction of the theory *for the time being*. [Zermelo, 1930, p. 403, emphasis following [Incurvati, 2020]]

So Zermelo also allows that there *may* be future use for non-well-founded sets, even if he was fully aware of the well-foundedness presented by the iterative picture. Again, attributing the concept ITERATIVE SET to Zermelo's use of the term "set" attributes a very trivial conceptual mistake to him, at least insofar as the «strong iterative conception» is concerned.

More sense can be made of Cantor, Mirimanoff, and Zermelo (among others) if we allow that perhaps their conception of SET was not firmly the iterative one. Rather than saddle them with a fully precise conception of set that not yet emerged, we can instead view their thought as imprecise and indeterminate between various conceptions that had yet to be fully isolated.

What is the situation with the current era? As we'll see later (Chapter 9), I think we are also living in a time of similar conceptual indeterminacy. I do think though, that nowadays we work with some form of «iterative conception» in mathematics, at least insofar as mainstream set theory is concerned. (The rub will come in that I think that the «iterative conception» is naturally modified to a conception that is indeterminate.) This is the situation as reported by Kanamori:

> It is nowadays almost banal that Foundation is the one axiom unnecessary for the recasting of mathematics in set-theoretic terms, but the axiom is also the salient feature that distinguishes investigations specific to set theory as an

[11] Indeed, for the «strong iterative conception» , it's sometimes not even necessary to assume that the stages are well-founded. For example, it can be proved in Button's Level Theory that the stages are well-founded.

autonomous field of mathematics. Indeed, it can be fairly said that current set theory is at base the study of well-foundedness, the Cantorian well-ordering doctrines adapted to the Zermelian generative conception of sets. [Kanamori, 1996, p. 28]

So, it seems relatively clear that much of the mathematical community has moved to a world in which the «iterative conception» of SET is now dominant. We may even say that the concept now associated with our term "set" is now ITERATIVE SET. But how did this change come about?

Point 2. Part of the dominance of the «iterative conception» can be explained by 'inference to the best conception'. How might we explain the dominance of the «iterative conception» ? Such an answer is probably best left to historians and sociologists studying mathematical histories and cultures, and so I won't go into too much detail here. But it bears mentioning that there are certain salient desirable features that a conception of set can have, that might result in it prevailing over its competitors. [Incurvati, 2020] refers to this as pursuing a strategy of *inference to the best conception*. Studying these factors—what is a *good making feature* of some conception of SET or other—*is* a task that falls to philosophers.

This is where some of the virtues identified in Chapter 2 can be brought to bear. There are constraints on what we want out of our conception of SET and the theories thereby motivated. In this regard, the «iterative conception» and ZFC perform spectacularly well. It easily provides a **Generous Arena** by the well-known constructions and thereby covers **Shared Standard** too. It provides a solid **Theory of Infinity**, giving us the means to do ordinal and cardinal arithmetic (though we will press on this issue later). It gives a particular arena in which the **Limits of Thought** can be examined and ZFC provides a **Metamathematical Corral**. Because of the independent account of what the universe looks like under the «iterative conception» , we also get some measure of **Risk Assessment**. This is especially so when we think of models like the *constructible universe*. L satisfies ZFC, but we also know a huge amount of what goes on in L (e.g. along with ZFC it satisfies CH among many other set-theoretic statements). This gives us confidence that if there were a contradiction lurking in ZFC, it would most likely turn up in L.[12] Moreover, L gives a very concrete picture of how the universe is formed under a very definite operation of construction (namely taking *definable* as opposed to *arbitrary* powersets).

[12] As [Steel, 2014, p. 156] puts it "...a voluble witness with an inconsistent story is more likely to contradict himself than a reticent one."

This gives additional confidence in the existence of a structure satisfying the axioms, in a similar way that we have confidence in the consistency of theories of arithmetic because of the picture of building up the structure of the natural numbers under the successor operation.[13]

This emphasises a very important fact that we will return to:

Consistency of the Iterative Conception. Most probably, the «iterative conception» (including both the «strong iterative conception» and «weak iterative conception») of SET are all consistent conceptions (and hence the concepts STRONGLY ITERATIVE SET, WEAKLY ITERATIVE SET, and ITERATIVE SET are all consistent concepts).

Regarding the paradoxes, the «iterative conception» certainly provides a **Testing Ground for Paradox**, giving us structures with which we can work in understanding how the paradoxes function against a probably consistent backdrop. But it also goes further, in that it provides:

Paradox Diagnosis. The relevant conception of set should provide an explanation for what goes wrong with the paradoxes.

And the iterative conception performs well here too. Both the Russell and Universal class have members unbounded in the stages, and so never get collected into a set. In this way, the «iterative conception» provides an account of why it is that **Universality** is at fault and **Indefinite Extensibility** holds.

(**Note:** Some philosophers, notably [Linnebo, 2010] doubt the efficacy of this explanation. See [Soysal, 2020] and [Incurvati, 2020] for rebuttals of these arguments. All said, I think it is open whether the «iterative conception» provides an *autonomous* reason to think that there is no Russell/Universal set, but once one has *accepted* the «iterative conception» it provides a perfectly good explanation of why there are no such sets on its own terms.)

The situation with the «stratified conception» is complex, and since my focus is on the «iterative conception» we will not consider it in detail. Some remarks are in order though to at least understand *why* the «iterative conception» rose to prominence. [Rosser, 1953] showed how to do some basic mathematical work such as arithmetic and analysis in NF. However NF-like set theories can be hard to work with. One issue is that they do not interact well with much of mathematical practice,

[13] I am grateful to Philip Welch for some discussion of this point. Of course exactly how much arithmetic one gets will depend on the other resources one allows, but the generation of an infinite domain by iterating successors is certainly a good start.

for instance NF *disproves* the Axiom of Choice, and one must always keep track of stratification (though perhaps this is just sociological inexperience). Regarding **Generous Arena**, we don't necessarily get representatives for the usual objects of set theory. Moreover, for many NF-based set theories, they have no model with the standard natural numbers, and are thus doomed to get facts about arithmetic wrong. Further, because Cantor's Theorem fails, stratified theories like NF do not guarantee the existence of the real numbers. (NAB: Check this.) It gives a particular arena in which the **Limits of Thought** can be examined. **Risk Assessment** can be answered, but only through a detour through type theory. In general, there is not much of an autonomous (independent of type theory) conception of what the universe is like for the «stratified conception» (in the way that there is with the «iterative conception»).

Some of these criticisms may well be sociological, and more to do with the mathematical community's general lack of familiarity with the systems rather than any deep-seated problem. These are the kinds of considerations pointed to by mathematicians considering NF. All of [Russell, 1959], [Martin, 1970], and [Boolos, 1971] are pretty scathing about the prospects of the theory as giving a motivated explanation of the paradoxes. Some have thought that the «stratified conception» does not provide a good **Paradox Diagnosis**, merely providing a syntactic restriction that suffices to block the paradoxical reasoning, rather than an explanation that is independently motivated by a conception of what the sets are like. This latter complaint will depend a little on one's taste for type-theoretic considerations. More generally these considerations are not knock-down, but represent reasons that have been advanced as to why the «iterative conception» is to be preferred over the «stratified conception» .

NAB: The above paragraph seems weak and too deep into the weeds.

Point 3. There is still a place for the «stratified conception» . So for SET, the «iterative conception» has become dominant, at least within mainstream mathematical circles. In particular, it (arguably) performs somewhat better with respect to several foundational constraints. However, the «stratified conception» still has a place as a theory of collections of a certain kind. [Incurvati, 2020] thinks it is better to think of the «stratified conception» as providing a theory of OBJECTIFIED PROPERTY, but at the end of the day the kinds of things that the «stratified conception» talks about are extensional and objectual, and so it makes sense to think of it as a conception of SET. Not too much

hangs on this, however, and one role we might want to put the sets postulated by the «stratified conception» to is for them to serve as the semantic values of predicates. Here, we might want sets corresponding to certain kinds of well defined predicate, and for these purposes, we might worry less about integration with the rest of mathematical practice. So there is a general question of set-theoretic pluralism here—what is the right conception might vary according to the job at hand. I won't take a stand on this issue (we'll leave pluralism as an open question in Chapter 11) but it will be useful to identify that this is a salient option.

This is where much of the literature ends—the «iterative conception» became the dominant conception of SET and effectively solved the paradoxes. In the rest of the book, I want to challenge this idea and argue instead that we are *ourselves* at a similar kind of conceptual crossroads. To make this clear, we'll first need to explain some set-theoretic constructions before developing some problems of our own.

Chapter 6

A closer look at some set-theoretic constructions

In this chapter, I want to present some tools from set theory, but do so very informally. Details are available in a wide variety of mathematical texts, and I feel that full detail would just bog down the reader in a book of this kind. What I want to do is present the general ideas so that:

(1.) The reader who might not have encountered these constructions before comes away with an idea of how they work.

(2.) Some connections to other areas of the philosophical literature are clear, in particular absolute generality.

We'll consider two kinds of idea. 6.1 will explain the technique of *forcing* in set theory. This is a way of adding *subsets* of sets to models. 6.4 will explain the idea of what I call 'climbing'—a cluster of techniques where we view the universe as a set in an expanded universe.

6.1 Forcing: The rough idea

Let's start with *forcing*, we'll start by approaching it from the angle of *what problem was it initially trying to solve*? We'll then examine *what role has it now taken on*? Finally we'll discuss *how does it work*?

As many will already know, the technique of *forcing* was developed in trying to prove that the continuum hypothesis is independent from ZFC. Since we knew that CH was true in L, we needed to find a way of making a model of $\neg$CH (and could then infer by the Completeness Theorem that neither CH nor its negation followed from ZFC, assuming ZFC consistent). Since we also knew that 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. And this is just what Paul Cohen did in [Cohen, 1963].

The rough idea of forcing to break CH is the following. You start with a model of ZFC that satisfies CH. 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 *functions* compared to the different sets of *reals*—so much so that any set of reals is either bijectable with the natural numbers or all the reals. $\neg$CH by contrast, says that there are lots of reals as compared to functions—so many that one can find some set x of reals that outstrips the bijections between x and $\aleph_0$, but also $2^{\aleph_0}$ outstrips x. So, if we have a model of CH, we want to add reals to that model in such a way that we keep ZFC and we also don't add any annoying bijections that might conspire to keep CH true. And that's what Cohen showed was possible with forcing, assuming some sets with nice properties exist, you can add these reals to make $\neg$CH true.

If you haven't encountered forcing much before, I want the reader to now stop and pause for just one minute and think about how, given the above observations and a model of $\neg$CH, we can make CH true again. What kinds of set could we add to a model of $\neg$CH (and what would we have to preserve) in order to restore CH again?[1]

The answer is that we need to add *functions* that provide the relevant bijections between sets of reals and either $\aleph_0$ or $2^{\aleph_0}$ without adding reals and preserving ZFC. Again, Cohen showed that forcing lets you do this. What one can do is *collapse* the cardinals between $\aleph_0$ and $2^{\aleph_0}$ to $\aleph_0$. In the new model, CH is true, since there are now bijections between $\aleph_0$ and the old 'cardinals' between $2^{\aleph_0}$ and $\aleph_0$. In fact, one can prove that so long as you have the nice sets required, a model of ZFC has an extension adding no new reals that satisfies CH, and an extension collapsing no cardinals that satisfies $\neg$CH.

So CH turned out to be independent from ZFC, and crucially, the proof relied on some idea of 'adding sets'. But regarding independence, it turns out that forcing is *incredibly flexible*. As Joel-David Hamkins writes:

> Would you like to live in a universe where CH holds, but $\diamondsuit$ fails? Or where $2^{\aleph_n} = \aleph_{n+2}$ for every natural number n? Would you like to have rigid Suslin trees? Would you like every Aronszajn tree to be special? Do you want a weakly

[1] 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 $\aleph_0$, but (alas) this won't work.

> compact cardinal κ for which $\Diamond_\kappa(REG)$ fails? Set theorists build models to order. ([Hamkins, 2012], p. 417)

The details of these mathematical examples don't matter so much, what is important is to note that there's a *huge* variety of statements that are independent from ZFC, and forcing is the main tool we have for showing this independence. This is to the point where in some areas of set theory a ZFC result is now something of a *surprise*—the norm is simply to find the right sets you need to force, and then possibly force again or restrict down (you can always move to a *smaller* model *after* forcing too) in the new model to obtain an independence result.

6.2 What role has forcing now taken on?

Here's a further philosophically important wrinkle to the story, one that's been underexamined in the philosophical literature.[2] Forcing is standardly seen as a too for proving *independence* results. We start with a model of ZFC, construct a model of ZFC + ϕ (or $\neg\phi$) and hence show that ZFC + $\neg\phi$ (or ϕ) is consistent if ZFC is. However forcing has now become so much *more* than a tool for proving independence. It is also a tool for *proving theorems* and *formulating axioms* about the universe.

The idea with proving theorems is the following: We know that certain properties in the universe will transfer to all forcing extensions, and vice versa. For example, if a set is countable in a model, then it will be countable in all larger models (since you can't kill off a bijection by *adding* sets). So, as long as we have suitable properties, we can move back and forwards between a universe and its forcing extensions. We may well have a bunch more information in a forcing extension (e.g. we can arrange things so we *know* that one of CH or ¬CH is true). Forcing then provides an 'information rich' context in which we can either find a set with such a suitable 'niceness property' that we can pull back to the universe, or if we're trying to perform a reductio, sometimes we can more easily find contradictions in the extension. Many of these examples are quite technical, so I won't give more details here (one can find some exposition in [Barton, 2020] and the references contained within), however often major open questions are solved using this methods.[3]

[2] Though I have examined it with others in a couple of articles [Barton, 2020] and [Antos et al., 2021].

[3] A recent example is [Malliaris and Shelah, 2016]'s surprising result that two uncountable infinities $\mathfrak{p}$ and $\mathfrak{t}$, which were thought to be separable, were in fact equal.

NAB: Is there a **super easy** example here?

A second way in which forcing is used, aside from proving theorems *within* ZFC, is to formulate *new* axioms. Let's see a couple of examples.

First we need to talk about *embeddings*. These provide a powerful way of postulating large cardinal properties for sets. An oft-used example from set theory is:

Definition 21. (ZFC) A cardinal κ is *measurable* iff there is a non-trivial elementary embedding j from V to an inner model M (that means that j that is not the identity map, and if V satisfies $\phi(x)$ then M satisfies $\phi(j(x))$—i.e. j preserves first-order truth) with κ the least ordinal moved by j.

Measurable cardinals are very strong, and have many inaccessible cardinals below them (for example, if κ is measurable, then there are κ-many inaccessibles below κ). But embeddings can be used with forcing to formulate interesting properties of cardinals:

Definition 22. (ZFC) A cardinal κ is *virtually rank-into-rank* iff for some $\lambda > \kappa$, there is a non-trivial elementary embedding $j : V_\lambda \to V_\lambda$ with κ the least ordinal moved by j.

That's a wee bit of mathematical notation, but the philosophical point can be summed up as follows: By postulating properties of cardinal numbers *in a forcing extension*, we can come up with new kinds of large cardinal axiom.[4] But this isn't a straightforward independence result, it's just a way of axiomatising a property that can be had by a particular cardinal number.

A different kind of example are principles of *absoluteness*, and these will be relevant for the rest of the book. Here's an example:

Definition 23. [Bagaria, 2000] *Bounded Proper Forcing Axiom (Absoluteness Version).* We say that V satisfies the *Absolute Bounded Proper Forcing Axiom* iff whenever ϕ is a sentence with only a single bank of existential quantifiers before a quantifier free sentence ψ, and ψ only has a certain class of restricted parameters (for the cognoscenti—from $H(\omega_2)$) if ϕ holds a forcing extension obtained by proper forcing (this is a particular restricted class of forcings) then ϕ holds in V.

The Absolute Bounded Proper Forcing Axiom gives us additional information about the kinds of sets that exist in V, by saying that if

[4] A survey is available in [Gitman and Schindler, 2018].

some kind of set exists in an extension, it already exists in the universe. Again here forcing is not being used to tell us *directly* about independence or relative consistency, but is rather giving us the expressive resources to make existential claims.

We'll discuss the significance of these claims for metaphysics in a moment (§6.6). First though I want to provide a little more mathematical detail on forcing for the interested reader.

6.3 Appendix A: A little more depth on forcing

NAB: This section is still very rough-and-ready, and I don't think I've quite pinned it down. It may just be better to delete it!

So, forcing is both a flexible tool for proving independence results, but can also be leveraged in proving theorems and formulating axioms. In this subsection I add a little more mathematical detail and provide an intuitive characterisation of forcing. I don't want the reader to get too bogged down, and this will include a little more mathematics than the previous sections, so it can be skimmed. Still, you'll get a better feel for what's going on later in the book if you do digest it (especially the notion of a *generic*) so I at least advocate giving it a go.

Before we go on, I should make the following clear: **There is more than one way of doing forcing, and I've opted for a way that appeals to me.** In particular, I've opted for the 'addition of sets to a model' approach (my terminology). But there's other options (e.g. algebraic approaches, such as those that use Boolean-valued models). Mathematically speaking, these are all equivalent for set forcing, and so which you use is something of a matter of taste. However I do grant that different approaches are differently philosophically suggestive (the Boolean-valued model approach doesn't press the issue with absolute generality so strongly). My aim here is to give the reader a feel for what I found to be the most understandable approach (and the one that presses my philosophical points here more acutely) but it's definitely not a complete story.

The way I suggest thinking of forcing is as a way of talking about 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. 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' for adding sets that we make take. (ii) $\mathbb{P}$-names, these are sets that can change their membership depending on what decisions we take from $\mathbb{P}$, (iii) *dense sets*, these are like *consultants*, 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 piece of advice given by a consultant. Let's look at these in more detail.

First, we need the notion of a *forcing partial order*:

Definition 24. A forcing partial order $\mathbb{P} = (P, \leq_{\mathbb{P}})$ is a partial order $\mathbb{P}$ such that:

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

(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 $s \leq_{\mathbb{P}} r$).

(iii) $\mathbb{P}$ is *separative*—for all $p, q \in P$, if $p \not\leq_{\mathbb{P}} q$ then there exists an $r \leq_{\mathbb{P}} p$ that is incompatible with q.

The way I'm going to suggest one thinks about this partial order is as an information space of *possible decisions* for adding in sets. As we'll see, we can define a class of sets that we can think of as 'gaining members' as we take decisions through $\mathbb{P}$. The conditions of being atomless and separative one can think of as conditions 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'.

The next idea we need is that of a $\mathbb{P}$-name. The definition looks somewhat complicated, but it can be given an intuitive backing.

Definition 25. A $\mathbb{P}$-*name* is a relation τ such that $\forall \langle \sigma, p \rangle \in \tau[$"$\sigma$ is a $\mathbb{P}$-name" $\wedge p \in \mathbb{P}]$. In other words, τ is a collection of ordered pairs, where the first element of each pair is a $\mathbb{P}$-name and the second is some condition in $\mathbb{P}$

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 members of p.

The intuition to have in mind is that a $\mathbb{P}$-name is the name for a possible set. Depending on what 'decisions' we take from $\mathbb{P}$ (we'll talk about this idea of 'decisions' in a sec, the key notion is that of a *generic* *(filter)*) we're going to rule in or out other $\mathbb{P}$-names (which in turn will be evaluated according to different decisions). So $\mathbb{P}$-names are kind of 'variable sets'—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 26. 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 *consultant*. 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 27. $G \subseteq \mathbb{P}$ is a *filter on* $\mathbb{P}$ iff:

(i) G is non-empty.

(ii) $p \in G$ and $q \geq p$ 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 consultant (i.e. dense set) in at least one place.

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

Definition 28. 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 set' 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 set'. When we give some G to the $\mathbb{P}$-names, we evaluate stepwise by analysing the valuation of all the names in τ and then either adding them to τ_G (if there is a $p \in G$ and $\langle \sigma, p \rangle \in \tau$) or discarding them (if there is no such $p \in G$). 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 the second coordinate is in G.

Here's an example, let's suppose we've got some $\mathbb{P}$. Consider the following $\mathbb{P}$-names.

$\tau = \emptyset$

$\sigma = \{\langle \tau, p \rangle\}$

$\mu = \{\langle \tau, p \rangle, \langle \sigma, q \rangle\}$

$\nu = \{\langle \tau, p \rangle, \langle \tau, q \rangle, \langle \sigma, p \rangle, \langle \sigma, q \rangle, \langle \mu, p \rangle, \langle \mu, q \rangle\}$

Let's suppose that $p \in G$ but $q \notin G$. What happens to our $\mathbb{P}$-names under G? Well, τ is trivial and so remains unchanged. We now have a value for τ, so σ, μ, and ν will contain $\tau_G = \emptyset$ (since we have $\langle \tau, p \rangle \in \sigma, \mu, \nu$). We now also know that $\sigma_G = \{\emptyset\}$, since we've evaluated all the names in σ. For μ, since we know $q \notin G$, 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 q (and so the evaluation of those names don't make it in via any ordered pair of the form $\langle \xi, q \rangle$) we also have that ν contains $\langle \tau, p \rangle$, $\langle \sigma, p \rangle$, and $\langle \mu, p \rangle$ 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 are like variable sets and 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 'consultants' (the dense sets) each of which can always present to you a way of continuing. 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 a decisions from $\mathbb{P}$. G then tells each $\mathbb{P}$-name who they are by ruling in and throwing out (evaluations of) $\mathbb{P}$-names based on what we were given by G.

So that's the rough idea of forcing. There's two points we should note. First:

Fact 29. If P is a forcing partial order in M, and G is $\mathbb{P}$-M-generic for $\mathbb{P}$, then $G \notin M$. In particular $\{p | p \in \mathbb{P} \wedge p \notin G\}$ 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[6]).

Fact 30. 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.[7]

This fact looks a little technical, but here's why it's 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 shows a 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 this was partly the inspiration for forcing, as reported by Cohen.[8]

6.4 Climbing: The rough idea

The rough idea of *climbing* is to turn classes of V into sets, rather than adding subsets to sets already in V. We can think, for example, of moving from some level V_α of the cumulative hierarchy to level $V_{\alpha+1}$. Now, the thought of climbing is that we might be able to view Ord as just another ordinal indexing a level, and move to levels V_{Ord+1} etc. In contrast to forcing which adds sets to some V_α for $\alpha < Ord$, climbing adds sets 'above' V_{Ord}.

The term 'climbing' does not refer to a specific technique (unlike forcing) and is perhaps less ubiquitous in set theory. There are, however, a cluster of techniques that are useful. Again, this can be both for proving theorems and formulating axioms. We'll consider just

[6] 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.

[7] Provide a [Kunen, 2013] citation for this—fixfix.

[8] Provide citation. fixfix.

two philosophically interesting ideas here, some others are available in [Antos et al., 2021].

The idea of climbing is most easily discussed in the context of *higher-order* set theory. Since we want to think of adding new sets 'on top' of V, it's helpful to introduce variables for *classes*, extensional entities that may not be sets. (**Note:** The philosophical question of how to interpret such talk is *super*-interesting, but I don't have space to discuss it here.[9]) One can then formulate theories like the following:

Definition 31. Gödel-Bernays set theory with a choice GBC, is the two-sorted theory that has, in addition to set-variables x, y, z etc., variables for *classes* X, Y, Z etc. As axioms it has (again, I opt for informal statements where possible):

(A) Set Axioms: The axioms of ZFC

(B) Class Axioms:

  (i) Classes with the same members are identical.
  (ii) Foundation: Every non-empty class has an $\in$-minimal element.
  (iii) Class Replacement: If F is a (possibly proper-class-sized) function, and x is a set, then the range of F on x is a set.
  (iv) Global Choice: There is a class that well-orders V.
  (v) Scheme of Predicative Class Comprehension: If ϕ is a formula of second-order set theory which may contain both set and class parameters and in which all quantification is restricted to sets, then there is a class containing all and only the ϕs.[10]

  In order to get the stronger Kelley-Morse set theory KM, one can strengthen this last axiom to:

  (vi) Scheme of Impredicative Class Comprehension: If ϕ is a formula of second-order set theory (which may contain both set and class parameters) and in which unrestricted quantification over classes and sets is allowed, then there is a class of all ϕs.

---

[9] The interested reader is directed to [Uzquiano, 2003] and [Welch and Horsten, 2016] for two papers providing positive accounts. A survey of some options is available in [Barton and Williams, U] (esp. §1)

[10] Here I suppress issues about ensuring that certain variables are free in the predicative and impredicative comprehension schemes in order to make the definitions more readable.

We can now discuss a couple of instances of climbing. The first is the addition of *truth predicates*. Presumably, the use of a truth predicate is desirable for talking about the universe. But as it turns out, building truth predicates for the universe is equivalent (over GBC) to building the constructible hierarchy starting with any class you like.[11]

The second is the use of possible extensions in formulating *large cardinal* axioms. This was examined by Reinhardt in [Reinhardt, 1974] and [Reinhardt, 1980]. Reinhardt's papers are notoriously difficult (e.g. [Reinhardt, 1980] uses 45 (!) axioms). A much cleaner presentation is available in Sam Roberts' thesis [Roberts, 2016] (see especially Chapter 3).[12] The rough idea again uses embeddings. We introduce modal operators $\Diamond$, $\Box$, and $@$. $\Diamond\phi$ should be thought of as saying that it is possible to add ranks to V to make ϕ true, and $\Box$ should be thought of as saying that no matter how one adds ranks to V, ϕ will be true, and $@$ is an operator that brings us back to the actual world. One can then formulate the following absoluteness idea (where $\vec{x}$ is some list of parameters):

Definition 33. $(\mathsf{R})\forall\vec{x}\big(\Diamond(\exists y)\phi \to (\exists y)\phi\big)$

This says that if there could have been a set such that ϕ, then there is a set such that ϕ. Roberts shows in his thesis that, in combination with some other modal axioms, R leads to very strong large cardinal axioms (known as 1-extendible cardinals—far stronger than a measurable cardinal).

6.5 Appendix B: A little more depth on climbing

Since climbing is a more general notion than forcing—it's a rough-and-ready informal idea that clusters several techniques together, rather

[11] This folklore result is presented very nicely in Kameryn Williams' thesis [Williams, 2018]. In particular Ch. 3, Theorem 3.14 (p. 123) shows that the following are equivalent:

Theorem 32. The following are equivalent over GBC class theory:

1. Elementary transfinite recursion for class well-orders.
2. For any class A and any well-order Γ, $Tr_\Gamma(A)$ exists (i.e. Γ-length iterated truth using A as a parameter).
3. For any class A and any well-order Γ, there is a code for $L_\Gamma(A)$.

[12] I am also grateful to Sam Roberts for some discussion of these ideas in conversation.

than a mathematically defined construction—I won't go into as much detail. There's two main ways that it pops up, and I'll just mention some of the rough intuitions. This is more for the curious reader; nothing later in the book will depend on this, so the reader who's had enough mathematical detail for now can safely skip it.

The first idea is just to handle things *axiomatically*. Here we introduce axioms that we take to govern the modal operators, any additional predicates we might want to have (e.g. a predicate $\bar{j}$ for an embedding j), and simply analyse what follows from the axioms. Whether or not these axioms are backed up by 'real' extensions is an important philosophical question, but not needed for studying what follows from the axioms. Since we've already discussed this earlier with Reinhardt's axioms, I won't go into any detail and will simply direct the reader to the clean exposition in Sam Robert's thesis [Roberts, 2016, Ch. 3].

The interrelation between *classes* and climbing deserves a mention though. The key point is that you can code sets by certain kinds of tree. In particular, you can think of a tree as coding a membership structure, you have the set you want to code at the top, all members on the level below, all elements of members on the next level down, and so on. (Effectively what you are doing here is coding up the structure of the transitive closure of $\{x\}$, for your desired set x.) But now, there's plenty of sets 'above' the universe we can code up using this technique. For instance, if we just replace every set in V by its singleton, we could then place all these singletons in a tree structure with the empty set as the top node to get a tree structure coding V as a set. This code is a very big tree (it's a proper class), but so long as we have a decent class theory we can talk about these codes and hence have a meaningful way of interpreting talk about 'sets' 'containing' proper classes using our class theory. This is an old folklore technique (it goes back at least to [Barwise, 1975]), but has found new application recently in analysing what can be done in class theory. See especially Carolin Antos' and Kameryn Williams' theses [Antos, 2015], [Williams, 2018], the resulting papers [Antos and Friedman, 2017] and [Williams, 2019]. There's also an interesting relationship between these tree codes, infinitary logics, and coding up extensions of the universe, [Antos et al., 2021] for the gory details.

NAB: More here, diagrams etc?

6.6 Connection to metaphysics and absolute generality

Before moving on, I want to bring together a few ideas that have occurred above, and connect them to a debate in philosophy.

Both climbing and forcing relate to the debate on *absolute generality*. In particular, do we think it is possible to quantify over all sets? Each of climbing and forcing challenge this assumption. This is because if you think that we can either climb or force over any universe, then any universe you pick will always miss out some sets—namely the relevant sets added.

Some think that there is great intuitive pull to these ideas. Regarding climbing, for example, 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?[13] [Linnebo, 2010, p. 147]

Linnebo then supports a principle that he calls **Collapse**—the principle that any things whatsoever form a set. The threat of paradox is avoided by holding that the range of the first-order quantifiers is modally indeterminate—given any definite domain we ascribe to them, there is a larger domain in which every class is formed as a set by lengthening the iteration of the powerset operation.

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:

> Like Galileo, peering through his telescope at the moons of Jupiter and inferring the existence of other worlds, catching a glimpse of what it would be like to live on them, set theorists have seen via forcing that divergent concepts of set lead to new set-theoretic worlds, extending our previous universe, and many are now busy studying what it would be like to live in them. [Hamkins, 2012, p. 425]

Moreover, as we've seen in this Chapter, there are various kinds of axioms and proofs that use apparent talk of extensions of 'the' universe. In this way, viewing any universe as just one among many and

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

situated within a wider network of universes allows us to interpret this talk very naturally. The situation is somewhat similar to how we can learn about various different geometries by modelling them in Euclidean space (an analogy pointed to by [Hamkins, 2012]) or how physicists can learn about certain models in general relativity by embedding them in larger ones, or via the universe as embedded within a multiverse, or the quantum mechanical state of the world by embedding it in a wider space.[14]

NAB: Probably we want more detail on these physical points right?

The use of these resources is highly suggestive of a kind of *set-theoretic* version of the indispensability argument. As many of philosophers of mathematics know, the indispensability argument proceeds roughly as follows (since there's so much literature here, I'll just write down a basic version of the argument):

Premise 1. We should accept the existence of those entities that are indispensable for our best scientific theories.

Premise 2. Mathematical entities are indispensable for our best scientific theories.

Conclusion. We should accept the existence of mathematical entities.

Going into the details of this argument would (obviously) take us too far afield. I merely wish to note that there's analogous one could make for climbing and forcing:

Premise 1. We should only accept the existence of entities that are indispensable for our best scientific theories (including set theory).

Premise 2. Climbing/forcing extensions are indispensable to our best set theories.

Conclusion. We should accept the existence of climbing/forcing extensions.

This argument shows that the mathematical *usefulness* of climbing/forcing extensions puts pressure on the generality absolutist *beyond* the mere 'intuitive plausibility' of the extending constructions.

[14] I thank Beau Madison-Mount and Neil Dewar for suggesting the example of general relativity, Don Howard for the (physicists') multiverse idea, and Rupert McCallum, Robert Black, and Colin McLarty for the suggestion of quantum mechanics.

Interestingly, both the standard and set-theoretic indispensability arguments have similar responses via a *hard road nominalism* where we try and provide a good interpretation of talk involving the relevant mathematical entities. (Note that the idea that there is a single maximal universe can be viewed as a kind of nominalism—nominalism about extensions of the universe.) There's a huge literature on 'standard' nominalism, starting with Field[15] and continuing to the present day. In set theory, a similar strategy is available—climbing and forcing extensions can be coded using sets from V. I won't say too much about this since it all gets very mathematically fiddly, but the rough idea is that one can find representatives for talk about these objects within any particular universe.[16] As an analogy, suppose you didn't believe that complex numbers existed, but only believed in the real numbers. You can still talk about complex numbers as *pairs* of real numbers, with the first coordinate representing the real part and the second the imaginary part (one then needs to go through and redefine the usual operations we want too, e.g. addition). So talk of real numbers (and pairs thereof) can *code* talk of complex numbers, without ever committing to the existence of complex numbers. So with climbing and forcing, though in a more complicated fashion—a universe V can internally 'code up' what would happen in an extension by talking about internal relations within V. Since the details are complicated, I'll say no more about it, but a good mathematical exposition for forcing is available in [Kunen, 2013], §IV.5, and some recent philosophical remarks of my own are available in [Barton, 2020] and [Antos et al., 2021] (this latter paper also discusses coding climbing extensions).

For indispensability, the question of course is whether these interpretations count as sufficiently 'natural'.[17] For now, let's just note that something always needs to be given up—the most 'natural' interpretation (putting aside worries about absolute generality) is to simply let the extensions be real. As Hamkins writes:

> ...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... 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

[15]See especially [Field, 1980] and [Field, 1989].

[16]Some of these ideas do need higher-order resources, others don't. See [Barton, 2020] for the case of forcing, [Antos et al., 2021] for the details of how to get a similar effect with climbing extensions.

[17]Again see [Barton, 2020] and [Antos et al., 2021] for discussion of this issue.

> 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]

There is a substantial debate as to how successful the coding techniques I mentioned are, and whether Hamkins' worry can be adequately answered. Even if we don't get into the guts of these coding constructions, philosophically speaking it's useful to isolate the following philosophical idea:

Capture. There is a universe conforming to the «iterative conception» of SET that is (i) unique (i.e. there's just one), (ii) maximal (i.e. it contains all possible sets), and (iii) cannot be extended.

Clearly **Capture** is important for not just the philosophy of mathematics, but philosophy more widely—it is essentially an affirmation of the ability to quantify over all the sets. Since so much of the literature on absolute generality rejects it because of the Russell/Cantor reasoning (i.e. often the case for a rejection of absolute generality is based on the specific problem of the *sets*), an acceptance of **Capture** will undercut many of the main arguments in this field. We should therefore keep it (and the interaction between **Capture** and climbing/forcing) in mind going forward—the question of whether **Capture** holds lies right at the intersection of the philosophy of set theory and metaphysics.

Chapter 7

The absoluteness conception of maximal iterative set

We're now in a position where we have the fundamentals of the «iterative conception» of SET down, understand ZFC, but also understand how there are set-theoretic constructions that bear on the issue of absolute generality. However in amongst all this, we saw that, via the definition of the V_αs (i.e. the cumulative hierarchy) within ZFC, we could say that the «strong iterative conception» was, in all likelihood, a consistent conception. We've also seen some of the mathematical constructions associated with set theory, and in particular the idea of *extending* universes by forcing or climbing. In this chapter, I want to:

(1.) Argue that the «iterative conception» is nonetheless *defective*.

(2.) Argue that there is, and has been for a while, a conception of ITERATIVE SET latent in set-theoretic practice, namely the «maximalist conception» .

(3.) Suggest that the «maximalist conception» is not sufficiently precise, but nonetheless has become an important conception in set theory, and accordingly many set theorists associate the concept MAXIMAL ITERATIVE SET with their use of the term "set".

(4.) Argue that a natural conception of MAXIMAL ITERATIVE SET is given by the «absoluteness conception» , on which any set that could exist, does exist.

So, let's get going!

7.1 The iterative conception of set is defective

The «iterative conception» seems to be consistent, and as I argued in Chapter 2 is a solid concept of set along numerous dimensions, not to mention that it is (in all likelihood) consistent. So why is it defective?

The core problem concerns **Theory of Infinity**. Recall that we want our conception of set to provide us with a rich and informative theory of infinite size. Whilst the «iterative conception» (axiomatised by ZFC) will provide me with all the tools I need to ensure that solutions to equations of ordinal and cardinal arithmetic *have answers*, it tells us *almost nothing* about the *values* of those solutions. Whilst there are some exceptions, generally speaking independence has been the norm. In particular, almost *any* pattern of the continuum function $f(\aleph_\alpha) = 2^{\aleph_\alpha}$ is consistent with ZFC (subject to some mild constraints[1]).

One might push this point further. The bare idea of iteration (one might think) does not get us very far at all. Tim Button has argued, for example, that a theory he calls *Level Theory* captures the bare idea of a process of iterative set formation, and this theory does not even imply that there are any sets (cf. [Button, 2021])! To get ZFC, we need the Axiom of Replacement to tell us that the ranges of certain functions exist. We also need the Powerset Axiom, which tells us that ever larger sets exist. These together (along with Infinity) are the real engine that gets set theory, ZFC, and the V_α hierarchy going.[2] So in order to get a **Generous Arena** you might think that we *need* something more than the mere concept of ITERATIVE SET.

Thus, while the «iterative conception» is *consistent*, it is *defective* in other ways—it fails to give us a sufficiently informative theory about its intended subject matter. What to do about this state of affairs?

7.2 Maximality and the sets

One move is to say that some kind of *maximality* is part of our notion of ITERATIVE SET (or should be added to it). This has been suggested by many authors. Famous oft-quoted passages include Hao Wang:

> We believe that the collection of all ordinals is very 'long' and each power set (of an infinite set) is very 'thick'. Hence

[1]For example König's Theorem tells us that $\kappa < \kappa^{cf(\kappa)}$ and is provable in ZFC.

[2]We should also note that the Axiom Scheme of Replacement and Infinity are equivalent to the Lévy-Montague reflection principle. Since I don't want to get into the details of reflection principles right now, I'll set this aside.

> any axioms to such effect are in accordance with our intuitive concept. [Wang, 1984, p. 553]

Penelope Maddy also advocates that some sort of richness should be part of set theory:[3]

> Contemporary pure mathematics is a vastly broad-ranging inquiry, dedicated to the notion that mathematicians should be free to study all and any structures and theories that seem to them of sufficient mathematical interest. If set theory is to found such a discipline, it should not impose any limitations of its own: the set theoretic arena in which mathematics is to be modelled should be as generous as possible; the set theoretic axioms from which mathematical theorems are to be proved should be as powerful and fruitful as possible. [Maddy, 1998, p. 142]

Gödel, in a footnote to the 1964 version of his paper on the Continuum Hypothesis, writes:

> On the other hand, from an axiom in some sense opposite to this one [here Gödel means $V = L$], the negation of Cantor's conjecture could perhaps be derived. I am thinking of an axiom which (similar to Hilbert's completeness axiom in geometry) would state some maximum property of the system of all sets, whereas axiom A [i.e. $V = L$] states a minimum property. Note that only a maximum property would seem to harmonize with the concept of set... [Gödel, 1964, pp. 262–263, footnote 23]

Here Gödel is contrasting the axiom that *all* sets are constructible (i.e. $V = L$) with the idea that there should be some 'maximum' property applying to the sets. The L_α construction keeps a very tight control on what sets are allowed at successor stages, and one can prove that L is the least transitive model of ZFC containing all ordinals under inclusion. So Gödel is saying that an axiom *opposite* to this (i.e. have some maximality property) is what meshes best with SET.

I only half-agree with Gödel. The *bare idea* that sets are formed in stages does not by itself legislate against $V = L$. This is witnessed by the fact that the constructible hierarchy has a perfectly good representation of even the «strong iterative conception» .

[3] Although it should be noted that Maddy has since changed her mind somewhat (see [Maddy, 2011]), and even in [Maddy, 1998] opts for an approach very different from what we'll consider in this book.

However, what I think that the above quotations show (and examples can be multiplied) is that there may be better conceptions of SET beyond the mere formation of stages as given by the «iterative conception» . I think that we are, in fact, already moving towards what I'll call the «maximalist conception» of ITERATIVE SET. The «maximalist conception» adds the additional constitutive principle:

Maximality There should be as many sets as possible.[4]

Before we continue, we should make a couple of remarks. **Maximality** is a principle that is very strange. Whilst it is a principle that many have regarded as providing an attractive sharpening of ITERATIVE SET, it's rare in science that we tolerate such ontological profligacy. The norm is often *parsimony* rather than *profligacy*, and it's interesting that in set theory this is not just tolerated, but encouraged.

Second, **Maximality** is also *philosophically* strange, in that normally we do not allow existence claims to be constitutive for concepts. For example, a simple version of the ontological argument for the existence of God relies on identifying a constitutive principle for GOD that they necessarily exist. But a natural (Kantian) rejoinder is that existence is not the kind of thing that we can say is *constitutive* for a concept. There is a possibility here that the «maximalist conception» of ITERATIVE SET also suffers from this malady, but we'll set this aside.

It may well be right that **Maximality** has been present (in some form or other) in «iterative conception» even from its early development (maybe the «iterative conception» of SET was always the «maximalist conception» of ITERATIVE SET for the purposes of representing the thoughts of our intellectual ancestors). Since I want to mark this distinction, however, we will speak of the «maximalist conception» of ITERATIVE SET, and separate it out from the «iterative conception» of SET simpliciter.

Unfortunately, whilst the «maximalist conception» might rule out some axioms (e.g. $V = L$) it doesn't improve the situation much with respect to **Theory of Infinity**. There are an *enormous* variety of proposed axioms that might be reasonably termed 'maximality principles' and they do *not* agree on much (see [Incurvati, 2017] for a survey).[5] Al-

[4]This is loosely adapted from a 'meta-axiom' provided by [Bagaria, 2005]:

(Maximality) The more sets the axiom asserts to exist, the better.

I want a version that is not so clearly tied to axioms (we're after conceptions *motivating* axioms after all) so I've gone for a more metaphysical version that will fit better with what comes later.

[5]I also raise some worries about the precision of 'richness' being employed in justificatory debates in [Barton, 2016].

most any principle (beyond obvious violators like $V = L$) can be given a 'maximality' spin.

Here's an example that avoids a lot of the technical detail of some of the axioms in the literature. CH can be regarded as a 'maximality principle', since it postulates the existence of lots and lots of *sets of reals*. However, $\neg$CH can also be given a maximality-style justification, since it says there are lots and lots of *functions* between sets of reals, the natural numbers, and the reals themselves. Moreover, this behaviour is witnessed by the forcing construction, you can make universes of sets bigger and bigger, flipping CH on and off. So it's unclear how to resolve in one way or the other if you think there should be 'lots' of sets, and that's even for what we might think is the *simplest* test question (namely CH).

NAB: Here's a way suggested to me by Tim Button that I think must be wrong, but I'm struggling to think why. Choice has a very easy **Maximality** justification: If I've got a set x at some stage V_α, then all elements of members of x must exist before V_α. So how could a set containing one of each of these members not exist at α too?

There's a similar argument for CH. Suppose we made the reals really really big. Well now we're iterating powersets over what we have, and we start forming the functions, so we should expect the relevant collapsing functions to exist. How should we respond to this? I think something funny is going on with the 'combinatorial' idea with AC, that isn't so clear with CH (e.g. there's no clear way of just singling out individual members out as there is with AC), you need the additional constraint that they form a function with such-and-such properties, and these are non-trivial things to have in play (rather than just merely noting that a bunch of things exist, so a set of them should also exist).

I *do* think though (as the above quotations suggest) that the «maximalist conception» of ITERATIVE SET is attractive, and may even have been latent pretty much from the off. Moreover, these ideas continue to resonate in many foundational debates. Therefore I think that these days there's at least one way to go—we treat MAXIMAL ITERATIVE SET as a concept in need of sharpening via a conception.

7.3 The absoluteness conception of maximal iterative set

We need some way of making the idea of MAXIMAL ITERATIVE SET precise. I'm going to examine *just one* of a *myriad* of options here, namely the «absoluteness conception» of MAXIMAL ITERATIVE SET. On the one hand this is quite a specific conception, and there are *plenty* of other options out there.[6] Whilst I do think that the «absoluteness conception» is a very interesting conception of MAXIMAL ITERATIVE SET—and importantly should be *appealing to philosophers*—we should also flag that we're now restricting our focus quite substantially. Part of the reason to do so is to to see how deep the engineering rabbit hole can go in set theory. (We'll discuss this limitation of our study in Chapter 11.)

The «absoluteness conception» of MAXIMAL ITERATIVE SET adds the following constitutive principle:

Absoluteness. If there **could** be a set such that ϕ then there **is** a set such that ϕ.

There's lots to be said about this principle. Let's start by identifying the fact that several set theorists have seen it as appealing. Magidor writes:

> The intuitive motivation[7]...is that the universe of sets is as rich as possible, or at the slogan level:
>
> A set [whose] existence is possible and there is no clear obstruction to its existence [exists]... [Magidor, 2012, p. 15]

One immediate problem is to say what we mean by "possible". Bagaria has the following answer:

> To attain a more concrete and useful form of the Maximality criterion it will be convenient to think about maximality in terms of models. Namely, suppose V is the universe of all sets as given by ZFC, and think of V as being properly contained in an ideal larger universe W which also satisfies ZFC and contains, of course, some sets that do not belong

[6] I'm not, for example, going to say anything about the inner model programme and Ultimate-L, which are undoubtedly on the table, though their relationship to **Maximality** is unclear.

[7] Here Magidor is talking about *forcing* axioms like MA and PFA.

> to V—and it may even contain V itself as a set—and whose existence, therefore, cannot be proved in ZFC alone. Now the new axiom should imply that some of those sets existing in W already exist in V, i.e., that some existential statements that hold in W hold also in V. [Bagaria, 2005] (NAB: Find page number—I only have the preprint right now.)

So—in the parlance we're employing here—ϕ is possible if it can be obtained by either climbing or forcing (Magidor, as it happens, favours forcing). If you're a fan of **Capture** this will have to be given some paraphrase (e.g. by the coding techniques discussed in §6.6) but the motivation is clear—if you could dream up or imagine a set such that ϕ holds of it, then you should have such a set in the universe.

Especially for what comes later, it will be useful to separate out the climbing and forcing versions of absoluteness. So we have:

Climbing Absoluteness. If there is a climbing extension such that $\exists x\phi$, then $\exists y\phi$.

Forcing Absoluteness. If there is a forcing extension such that $\exists x\phi$, then $\exists y\phi$.

It should be noted that there's a few formal details that I'm suppressing for clarity. First, since climbing is not formally defined, **Climbing Absoluteness** really refers to a cluster of formally precise principles. **Forcing Absoluteness**, as we'll also see, can be controlled in various ways. So these should be viewed not as formal definitions, but rough ideas that can be used to *motivate* formally precise axioms. Really, one has to specify the base language, introduce a modal logic, and axiomatisation etc. Much of this work has already been done[8] and nothing in my arguments is going to depend on a particular formalisation. Giving one I feel would just muddy the (already murky) philosophical waters, so I'll just proceed informally and flag where needed any formal subtleties.

Second, if you're a fan of **Capture**, we'll need some paraphrase so as not to trivialise the principles (since if no extensions exist, then the implications hold vacuously). Thankfully there's ways of *coding up* talk of extensions (we discussed some of these in Chapter 6), and so even if you don't think there's *really* such things, we have ways of talking *as if* there are. The details are in [Antos et al., 2021], but I'll

[8] For climbing see [Roberts, 2016], [Linnebo, 2010], [Linnebo, 2013], and [Hamkins and Linnebo, 2022], and for forcing see [Hamkins and Loewe, 2008] and [Hamkins and Linnebo, 2022].

set this aside from now on. Suffice to say that even if you're a fan of **Capture** there's still ways for you to interpret "It's possible that there's a climbing/forcing extension with a set such that ϕ". From now on, I'll just bracket this and work informally.

Further, eagle-eyed readers may notice a problem with **Forcing Absoluteness** immediately: What if ϕ is CH? Didn't I just say that both CH and $\neg$CH were forceable? And both can be given an existential spin—CH says 'There is a set x that is the power set of the natural numbers, and every set of reals is either bijective with x or bijective with the natural numbers' and $\neg$CH says 'There is a set of reals x such that there is neither a bijection between the natural numbers and x nor the between the continuum and x'.[9]

A solution (and the one that I'll adopt here) is to move to absoluteness for existential claims that are otherwise *quantifier-free*.[10] A motivation for this can be seen from the fact that if you have an existential quantifier followed by a universal quantifier followed by some quantifier-free formula ϕ, its often possible to make $\forall y\phi$ true of some x by *getting rid* of some of the objects in the domain. However if we just had $\exists x\phi$ with ϕ quantifier free, we couldn't pull this trick. So from now on I'll just restrict to these nice sentences. Again though, these formal details can be suppressed for everything that follows.

7.4 Progress under the absoluteness conception

What axioms might the «absoluteness conception» motivate? Well, we've seen some in Chapter 6 that conform to this template. Let's revisit them.

First we had the *Absolute Bounded Proper Forcing Axiom* which asserted:

The Absolute Bounded Proper Forcing Axiom. V satisfies the *Absolute Bounded Proper Forcing Axiom* iff whenever ϕ is a sentence with only a single bank of existential quantifiers before a quantifier free sentence ψ, and ψ only has parameters from $H(\omega_2)$, if ϕ holds a forcing extension $V[G]$ obtained by proper forcing (this is a particular restricted class of forcings) then ϕ holds in V.

Putting aside all the technical jargon about parameters etc., the Absolute Bounded Proper Forcing Axiom asserts **Forcing Absoluteness**

[9] In set-theoretic parlance, both CH and $\neg$CH are Σ_2-sentences.

[10] In set-theoretic parlance, we restrict to Σ_1-formulas.

for many formulas of the kind we suggested above. And, as a bonus, the Absolute Bounded Forcing Axiom implies ¬CH.

Second, we had principle R:

(R) $\forall \vec{x}\big(\Diamond(\exists y)\phi \rightarrow (\exists y)\phi\big)$

Principle R, suitably formulated (here specifics about the modal theory do matter, we need an actuality operator to get significant strength), yields an enormous variety of *very* strong large cardinals.

So, this looks promising! Black-boxing the mathematical details, the «absoluteness conception» appears to be making some progress—we have a conception of MAXIMAL ITERATIVE SET that implies that there are lots of large cardinals and that CH is false. This is not a panacea, and there are still open questions regarding how to motivate axioms that do more under the «absoluteness conception» . (For example, it's not clear how to get traction on the continuum function beyond $2^{\aleph_0}$.) However it seems like a good start. But is there a problem lurking in the shadows?

Chapter 8

A 'new' kind of paradox?

In this chapter I want to argue that the «absoluteness conception» of MAXIMAL ITERATIVE SET is in fact inconsistent. Just to foreshadow, in the next chapter, I'll then argue that this is, in fact, a situation that many set theorists understand well, and some proposals can be seen as advocating conceptual engineering of a similar variety to what we saw with (i) TRUTH, ASCENDING TRUTH, and DESCENDING TRUTH, and (ii) the «naive conception» ,«stratified conception» , and «iterative conception» . In particular we have constitutive principles of an inconsistent concept that can be traded off against one another.

8.1 The Cohen-Scott Paradox

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. However, recent work has developed these ideas substantially, including work by [Meadows, 2015], [Scambler, 2021], [Builes and Wilson, 2022], and [Barton and Friedman, U].[1]

Let's briefly drop the Powerset Axiom (we'll see why in a moment), but keep **Climbing Absoluteness** and **Forcing Absoluteness**. What should we say on this picture about the existence of uncountable sets and large cardinals? Well, it's very plausible to think that it's possible that there are uncountable sets. If we climb and make the ordinals of V a set, for instance, then this set should be uncountable. So, by **Climbing Absoluteness**, there's at least one uncountable set x. But now it seems possible that we could climb again, make the ordinals of our universe Ord an uncountable set with a smaller uncountable

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

set below (i.e. x). By **Climbing Absoluteness**, there should be *two* uncountable sets in V. More generally, for any ordinal α, there should be an uncountable set larger than α simply because by climbing we can *make* one (namely the ordinals). So **Climbing Absoluteness** pushes us in the direction of many uncountable sets (in fact a proper class of them).

Similar remarks apply to any large cardinal axiom that you think holds of the ordinals. If you think that it's possible that the ordinals of our universe are an inaccessible cardinal in a climbing extension, then we can pull the same trick as we did before. In fact, it's a small step to getting a proper class of inaccessibles using the same reasoning as in the previous paragraph.

All this can be formalised within modal logic if so desired, but I won't do so here.[2] **Climbing Absoluteness**, in combination with suitable well motivated principles about what's possible, pushes us in the direction of uncountable sets, and stronger and stronger large cardinals. The key point is that the universe of sets exhibits a lot of *closure* properties—since the universe transcends any particular operations we'd care to give, there are sets that do so too.

What does **Width Absoluteness** tell us? Well, given any set x, we can force to collapse the cardinality of x to be countable. Moreover, as it turns out, the formula required to assert this is exactly one of the form that starts with an existential quantifier and is followed by a quantifier-free formula. (For an intuitive way to see this, note that you can't make a countable set x uncountable by adding more sets—once you've got a bijection from x to the natural numbers, you've still got one no matter how much stuff you add—set theorists often say that countability is *upwards absolute*.) So, using **Width Absoluteness**, we should think that x is in fact countable. Since the choice of x was arbitrary, we should in fact think that every set is countable.[3]

This fact that forcing seems to want to kill off uncountable sets was noticed by Scott in his introduction to Bell's book on forcing:

> 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

[2] One just needs the axioms $\Diamond\exists x\phi(x) \rightarrow \exists y\phi(y)$, and then the possibility axioms for whatever properties you think the ordinals can have (e.g. $\Diamond\exists x$ "x is inaccessible").

[3] In fact, this version of absoluteness (which we call the *Weak Axiom of Set Generic Absoluteness* in [Barton and Friedman, U]) is *equivalent* (over ZFC minus Powerset) to the claim that every set is countable. See Fact 16 of [Barton and Friedman, U].

> 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. But really pleasant axioms have not been produced by me or anyone else, and the suggestion remains speculation. A new idea (or point of view) is needed, and in the meantime all we can do is to study the great variety of models. [Scott, 1977, p. xv]

Of course, since this level of **Forcing Absoluteness** contradicts ZFC we might want to temper it slightly. There are ways of achieving this. For example the *Inner Model Hypothesis* is a kind of absoluteness principle that implies that there are no inaccessible cardinals in the universe, but is nonetheless (probably) consistent with ZFC. Since the details are a little mathematically involved, I'll relegate them to a footnote.[4]

What is going on here? The core point is the following: **Forcing Absoluteness** wants to *kill off* closure properties by saying that there are lots of subsets floating around, subsets that can then be used to code functions and witness the failure of closure properties. **Climbing Absoluteness**, on the other hand, wants to say we have *lots* of sets that look like the universe in various ways, and in particular possess

[4]One formulation of the Inner Model Hypothesis says that anything true in an inner model of an outer model of the universe is already true in an inner model of the universe. For us though, it's more interesting that we get these anti-large cardinal properties for a slightly weakened version. In particular *class forcing* is a kind of forcing where we allow that the forcing partial order can also be a proper class. A class forcing is *tame* if it preserves both Replacement and Powerset (the definition is technical to state, so I omit it). We start with the following definition

Definition 34. (GBC) A formula is *persistent*-Σ^1_1 iff it is of the following form:

$$(\exists M)(\text{“}M \text{ is a transitive class”} \wedge M \models \psi)$$

where ψ is first-order.

We can then define:

Definition 35. (NBG) *Tame parameter-free persistent* Σ^1_1*-absoluteness* is the claim that if ϕ is persistent-Σ^1_1 and true in a tame class-generic extension of V, then ϕ is true in V.

One can show that tame parameter-free persistent Σ^1_1-absoluteness implies that there are no inaccessible cardinals in V. See [Friedman, 2006] for the result, [Friedman et al., 2008] for some consistency implications, and [Barton, Sa] for some discussion of the philosophical implications for large cardinal axioms and formulating the principle in terms of tame class forcing. These principles are consistent relative to ZFC plus the existence of large cardinals.

a wide variety of closure properties. So we cannot have both **Forcing Absoluteness** and **Climbing Absoluteness** in full generality, and even restricted versions can come into conflict with one another. So without serious revision, the «absoluteness conception» of SET is inconsistent.

Before we continue, I want to emphasise: No reasonable set theorist has ever accepted both **Climbing Absoluteness** and **Forcing Absoluteness** in this generality. Set theorists are not dummies, and are able to see this contradiction coming a mile off. This is exactly why many absoluteness principles are formulated with a lot of caveats on (i) the complexity of the formulas allowed, (ii) the kinds of extension we consider, and (iii) the parameters we permit. The point behind the «absoluteness conception» is not that individual agents hold it, but (as we'll argue later) we are pulled as a community in different directions by its different constitutive principles. Some people choose to go one way, others a different way. We'll explore this in more detail shortly (in Chapter 9). For now I want to show how the Cohen-Scott Paradox is closely linked to another paradox of infinity—*the Enumerator Paradox*—before moving on to consider the relationship between the Cohen-Scott Paradox and diagonalisation.

8.2 The Enumerator Paradox

The Cohen-Scott Paradox is closely linked to another paradox put forward by [Meadows, 2015]. Meadows begins by noting that, in reference to Cantor's Theorem:

> The idea of there being multiple sizes of infinity is very strange and mysterious. [Meadows, 2015, p. 196]

To formalise the idea that all sets have the same infinite size, Meadows suggests the following axiom:

Definition 36. (INF) For all x, either x is finite or there is some y coding a bijection between y and the natural numbers.

By adding (INF) to our theory of sets, Cantor's Theorem can be seen as a proof that the resulting theory is inconsistent. Meadows calls this *the Enumerator Paradox*.

The connection between the Enumerator Paradox and the Cohen-Scott Paradox is that one can see the use of **Forcing Absoluteness** in the Cohen-Scott Paradox as a way of justifying (INF). If we do think that **Forcing Absoluteness** should be given its most general reading, then it naturally leads to (INF)—every set is going to be countable.

The point to be made here is just that although it might *seem* like the obvious thing to be thrown out is the level of **Forcing Absoluteness** advocated, there are intuitions (e.g. (INF)) that mesh well with every set being countable. Perhaps, instead, we should drop the Powerset Axiom, and work with a weakening of ZFC?[5]

8.3 The Cohen-Scott Paradox and Diagonalisation

This naturally leads us into the following question: What is the link between Russell's Paradox, Cantor's Theorem, and The Cohen-Scott Paradox?

We have already seen a tight link between Russell's Paradox and Cantor's Theorem in Chapter 4—in the case where we take the universal set, the identity surjection/injection and run the standard proof of Cantor's Theorem, we get the Russell set.

The observation now is to note that the non-existence of a bijection from the natural numbers (for brevity, we'll denote this by ω in this chapter) to an arbitrary set x can also be put into this form. Suppose that there is in fact a bijection $G : \omega \rightarrowtail\!\!\!\!\rightarrow x$, obtained by a collapse forcing $\mathbb{P}$.[6] 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 it in Chapter 6.

With this little bit of set-theoretic reasoning out of the way, we can then say that the assumption of such a generic in the universe leads to contradiction. Since G is generic and E is dense, we know that G intersects E at some point p. We then have:

$$p \notin G \Leftrightarrow p \in E \text{ (definition of } E\text{)}$$

$$\Leftrightarrow p \in G \text{ (since } G \text{ is dense and intersects } E \text{ at } p\text{)}$$

So the assumption of the existence of a generic (in particular one coding a bijection $f : x \rightarrowtail\!\!\!\!\rightarrow \omega$) leads to contradiction. But the point to note is that there is a similarity to the Russell/Cantor reasoning. There we had the assumption of the existence of a particular injection/surjection leading to contradictory claims about (non-)self-membership. Here we have the existence of an injection/surjection, whilst not leading to contradictory claims about self-membership, we do have the

[5] This is the *tentative* suggestion proposed in [Barton and Friedman, U].

[6] In fact we need only assume that x is countable, since that will allow us to produce the generic for $\mathbb{P}$, but I suppress this detail for clarity.

contradictory $p \in G \leftrightarrow p \notin G$. Thinking of $p \in G$ as the claim that p was in the 'good decisions' given by G, we have the result that p is a G-good decision iff it is not a G-good decision. So whilst the analogy is not perfect, we have a diagonal-style contradiction obtained by assuming the existence of a particular injection/surjection.

What these similarities show is that *if* one is tempted to reject **Capture** on the grounds of Russell's Paradox, one is the *also* pressed to reject **Capture** on the basis of the Cohen-Scott Paradox. To be sure, there are some disanalogies—there isn't the same self-reference in the Cohen-Scott contradiction as there is in the Russell-Cantor reasoning, and we need the combinatorial lemma that $\mathbb{P} - G$ is dense. However, if one wants to reject **Capture** on Russellian grounds, there is an explanation to be given as to why we should not also reject **Capture** for forcing too.

Chapter 9

Contemporary engineering

We now have a contradiction in the «absoluteness conception» of MAXIMAL ITERATIVE SET, which can be traced to a conflict between **Climbing Absoluteness** and **Forcing Absoluteness**. We've also seen how these kinds of questions are intimately related to **Capture**. In this chapter I want to argue that many programmes in the Philosophy of Set Theory can be re-examined in this light. In fact, by drawing an analogy between (1.) the case of ASCENDING TRUTH and DESCENDING TRUTH arising from TRUTH, (2.) the «iterative conception» and «stratified conception» arising out of the «naive conception» of SET, we can view many views as proposing two conceptions of set, the «climbing absoluteness conception» and «forcing absoluteness conception» of MAXIMAL ITERATIVE SET. I want to argue that much of contemporary foundations of set theory thus fits into the conceptual engineering mould. However I also want to argue that there is a separate *metaphysical* question lying behind many of these theory choices, namely the question of **Capture**.

9.1 Views on the ontology of set theory

Many views in the philosophy of set theory centre on the notion of **Capture** and the idea of **Multiversism**.

One has the following view:

Universism. Capture is true.

Universism is often contrasted with:

Multiversism. There is no one maximal universe of set theory, if we are given a universe of set theory, it can always be extended.

Multiversism implies the falsity of **Capture**, since there's no one universe, there's no one universe for the «iterative conception» to describe.

Multiversism is often subdivided into the following kinds:

Climbing Multiversism. Given a universe, there is a *climbing extension* of that universe in which it is a set.

Forcing Multiversism. Given a universe and a forcing partial order in that universe, there is a forcing extension of that universe containing a generic for that partial order.

Schematic multiversism. Any universe satisfying ZFC (and possibly weaker theories) is a legitimate universe of set theory, and there is no one maximal such model.

Often **Climbing Multiversism** and **Forcing Multiversism** are described as multiversisms about 'height' and 'width' respectively (the former because adding ranks adds ordinals and makes the universe 'taller', the latter because adding subsets to some stage makes that stage 'fatter'). **Schematic Multiversism** is perhaps a misnomer, it doesn't immediately refute **Universism**, but is usually combined with climbing- or forcing-multiversist principles (e.g. [Hamkins, 2012]).

One issue that I think is interesting, but I'll suppress for the sake of space, is the issue of whether the different kinds of **Multiversism** are all species of the following view:

Potentialism. *Potentialism* is the view in the philosophy of set theory that sets gradually come into being as part of a modal process.[1]

Potentialism has a long and rich history, versions of potentialism about arithmetic have been isolated even in the work of Aristotle. Any kind of **Multiversism** naturally leads to modal descriptions of the subject matter of set theory—we can think of the universes as worlds and (for instance) climbing and forcing as ways of moving between them. However, it's open whether or not any brand of multiversism is also a kind of potentialism, one might think that the latter demands some description of a process that is iterated, and this is not obviously present (especially with Schematic Multiversism). This perhaps has metaphysical implications, many potentialists, for example speak as though there is *one universe* but that it is *modally indefinite*, in contrast to multiversists who are clear that there are *multiple universes*. We'll

[1] This definition is adapted from [Hamkins and Linnebo, 2022].

suppress this detail, and just consider multiversism—though potentialist ideas will be in the background throughout.

Many combinations of this kind have been proposed. [Linnebo, 2010] and [Linnebo, 2013] are climbing multiversist/potentialist. [Steel, 2014] proposes a picture on which every universe can be forced over but not climbed over, an idea which is taken up by [Meadows, 2015]. [Scambler, 2021] is both climbing and forcing multiversist/potentialist, as is [Arrigoni and Friedman, 2013]. Finally [Hamkins, 2012] is a schematic multiversist who is also a forcing and climbing multiversist. Normally these views are viewed as *metaphysical* disagreements—is there/are there universes of such-and-such a kind—and are thus centred around the issue of **Capture**. What I want to now illustrate is that each can be viewed as stances on the **Forcing Absoluteness/Climbing Absoluteness** dichotomy, and thus as different ways of engineering the «absoluteness conception» of MAXIMAL ITERATIVE SET. These views, I argue cluster *conceptually* into those that *reject* one of **Climbing Absoluteness** or **Forcing Absoluteness**. We'll see that the end result are distinct conceptions of set that are likely consistent and provide their own answers to **Theory of Infinity**. **Capture**, I argue, is an important motivating question, but not necessarily related to what

9.2 Mirroring Theorems

Before we get into the details of how we should view contemporary programmes in terms of engineering, we need a discussion of *Mirroring Theorems*. These tell you how you can go between the modal theories proposed by the multiversist to non-modal theories. The modal theories in question normally employ operators $\Box$ and $\Diamond$, with $\Box\phi$ interpreted as "in all future universes ϕ" and $\Diamond\phi$ meaning "there is a future universe such that ϕ". A natural translation, if you're a forcing or climbing multiversist is to say that we should interpret the 'normal' set-theoretic quantifiers $\forall x$ and $\exists y$ as the modalised $\Box\forall x$ and $\Diamond\exists y$. Given a sentence ϕ in a 'normal' set theory, we can find its 'potentialist translation' $\phi^{\Diamond}$ by replacing quantifiers in this fashion. The point being that when we operate mathematically we just want to speak as if we're quantifying over sets as normal, and what we mean by this is that, in the potentialist/multiversist's sense, we *could* get (or *always* get) sets such that ϕ.

What we've noticed in the last couple of years is that as long as the underlying logic of the modal system is S4.2, we will have the following:

Theorem 37. $\mathsf{T}^{\Diamond}$ proves $\phi^{\Diamond}$ iff T proves ϕ.

This has been done in the case of Climbing Multiversism [Linnebo, 2010] to show that the modal theory interprets ZFC, and [Scambler, 2021] to show that a framework that is both forcing and climbing multiversist satisfies ZFC minus Powerset with the sentence that every set is countable added.

9.3 How each engineers

Let's start by recalling the case of TRUTH there we saw that the constitutive principles for TRUTH were inconsistent, and so (following [Scharp, 2013]) we can opt for a pair of replacement concepts, ASCENDING TRUTH and DESCENDING TRUTH, corresponding to each direction of the Tarski biconditionals. We saw that there was an analogy here with the «naive conception» of SET, which further split down into the «stratified conception» (accepting **Universality**, rejecting **Indefinite Extensibility**) and the «iterative conception» (accepting **Indefinite Extensibility** and rejecting **Universality**). We've now moved to the «absoluteness conception» of MAXIMAL ITERATIVE SET, and seen that **Forcing Absoluteness** and **Climbing Absoluteness** conflict with one another (when formalised in full generality). As the reader might be able to guess, we're going to identify two conceptual roads we might take:

The Climbing Absoluteness Conception. The «climbing absoluteness conception» of MAXIMAL ITERATIVE SET holds that **Climbing Absoluteness** should be given priority, and **Forcing Absoluteness** should be curtailed.

The Forcing Absoluteness Conception. The «forcing absoluteness conception» of MAXIMAL ITERATIVE SET holds that **Forcing Absoluteness** should be given priority, and **Climbing Absoluteness** should be curtailed.

Viewing the «absoluteness conception» in this light, it should be clear that just as in the case of TRUTH and the «naive conception», these are both legitimate conceptual moves that produce legitimate conceptions of MAXIMAL ITERATIVE SET. What I want to argue now is that many of these moves are already present in the literature on the philosophy of set theory.

9.4 The Climbing Absoluteness Conception

The Climbing Absoluteness conception has been advanced by several authors, and it would take too long to go into detail on everyone.[2] Bagaria, as a clear advocate of the «absoluteness conception» provides an excellent example. We've already noted that he considers climbing absoluteness as a legitimate kind of absoluteness (recall his talk of absoluteness between the universe and models containing the universe as a set).[3] However he also often suggests that there is just a single universe and that **Capture** holds. For example, he writes:

> Of course, there are no real extensions of the universe of all sets, and therefore no real forcing extensions. [Bagaria, 2005] (NAB: Find page number, I only have a pre-print right now.)

As a result though, Bagaria thinks that **Forcing Absoluteness** needs to be constrained. In particular he characterises it as follows:

> Generally speaking, an axiom of generic absoluteness asserts that whatever statement can be forced is true, subject only to the requirement that it be consistent. [Bagaria, 2005], pp. 21–22 of preprint NAB: Need to find correct pagination.

The key point to note here is that we need to ask *consistency with what*? Bagaria has in mind ZFC and large cardinals, writing:

> All known large-cardinal axioms are compatible with Bounded Forcing Axioms. Thus it is reasonable to work with both kinds of axioms simultaneously. [Bagaria, 2005], NAB: p. 25 of pre-print, find correct citation.

---

[2]Notable proponents include (time slices of) Joan Bagaria, Peter Koellner, Leon Horsten, Penelope Maddy, Donald Martin, Philip Welch, Hugh Woodin.

[3]I repeat the quotation for ease of reference:

> To attain a more concrete and useful form of the Maximality criterion it will be convenient to think about maximality in terms of models. Namely, suppose V is the universe of all sets as given by ZFC, and think of V as being properly contained in an ideal larger universe W which also satisfies ZFC and contains, of course, some sets that do not belong to V—and it may even contain V itself as a set—and whose existence, therefore, cannot be proved in ZFC alone. Now the new axiom should imply that some of those sets existing in W already exist in V, i.e., that some existential statements that hold in W hold also in V. [Bagaria, 2005]

But as we've mentioned, there are Forcing Absoluteness principles that kill off ZFC, and others that kill off the existence of inaccessibles. This means that there *are* choices to be made, just because Bagaria regards ZFC as an epistemological hinge does not mean that a conceptual choice has not been made (as witnessed by the curtailing of **Forcing Absoluteness**).

As we noted earlier, the idea of Climbing Absoluteness is suggested by denying **Capture** and moving to a picture on which any universe can be climbed over. This is the approach of Tait who, after proposing a principle that entails universes with climbing absoluteness properties and appreciatively explaining Zermelo's version of Climbing Multiversism, writes:

> But, contrary to a common understanding of set theory as a theory about 'the' universe of sets, there is no such universe: the notion of 'all ordinals' or 'all sets' lacks rigorous mathematical sense. [Tait, 2005], p. 11 of preprint, NAB: Find correct pagination.

In my own work, I've argued that the «climbing absoluteness conception» of MAXIMAL ITERATIVE SET might motivate stronger large cardinal axioms along the lines of the Universist, even if we are Climbing Multiversists. For example:

> If the Universist can appeal to a general principle of 'richness' that follows from an unfolding of our concept of set, then so can the [Climbing Multiversist]. This provides the material for the [Climbing Multiversist] to generate a response to the Universist. The [Climbing Multiversist] may assert that she agrees, on the basis of our (enriched) concept of set the subject matter of Set Theory should be as structurally rich as possible and contain many and varied sets. Given this, she can say, it is entirely reasonable that we should expect (as part of a conceptual analysis of this 'rich' conception of set) there to be a $\mathcal{V} = V_\beta$ that reflects properties to its initial segments. [Barton, 2016] NAB: Find correct pagination.

If this is the view proposed, then both the **Universist** and the **Climbing Multiversist** who endorse the «climbing absoluteness conception» of MAXIMAL ITERATIVE SET have at least partial answers to **Theory of Infinity**. **Forcing Absoluteness**, they might contend, is desirable so long as it doesn't conflict with **Climbing Absoluteness**. And as

Bagaria notes, many of the **Forcing Absoluteness** principles that imply ¬CH are consistent with large cardinals. So the advocate of the «climbing absoluteness conception» is pushed towards many large cardinals and ¬CH. And moreover, the mirroring theorems ensure that we can move backwards and forwards between these two pictures, whatever one's view on **Capture**.

Moreover, this conception is consistent with the «strong iterative conception» . All possible sets can be formed at successor stages (whether or not **Capture** holds).

These factors are relevant for the strategy of *inference to the best conception*. We have large cardinals, ¬CH, and the conception is *strongly iterative* (and hence also *weakly iterative*).

The major conceptual open question for the advocate of the «climbing absoluteness conception» is how to generalise **Forcing Absoluteness** in a manner consistent with large cardinals, and in such a way that other independent sentences (e.g. the value of the continuum function more generally) are settled.

9.5 The Forcing Absoluteness Conception

Instead of accepting **Climbing Absoluteness** we might instead accept **Forcing Absoluteness** in its full generality, and curtail **Climbing Absoluteness**. Again, we can examine those views that validate **Capture** and those that don't.

Proposals that validate **Capture** are a little thin on the ground, precisely because a wholehearted endorsement of **Forcing Absoluteness** in its widest possible sense is going to yield the immediate result that every set is countable. However, we might follow Scott here and accept that the continuum is a proper class and that there are still uncountably many real numbers, it is just that now "uncountable" also implies "proper-class-sized".

In my own work (with Sy-David Friedman) I have proposed a view that incorporates a strong forms of **Forcing Absoluteness**, enough to obtain the result that every set is countable (see [Barton and Friedman, U]). Consistency is obtained by getting rid of the Powerset Axiom. **Climbing Absoluteness** has to be substantially curtailed too—one can have first-order properties held by V in *transitive sets*, but higher-order properties are out, and even quite weak forms of **Climbing Absoluteness** have to be barred—even the generation of a single uncountable set will yield inconsistency.

Of course, a natural immediate question is: What becomes of ZFC in this context? The quick answer is that you can still have ZFC in *inner models*. You just have to leave out the subsets that witness bijections with the natural numbers. In fact, one of the axioms proposed in [Barton and Friedman, U] (the *Ordinal Inner Model Hypothesis*) implies that every set is countable but also that ZFC with large cardinals added holds in inner models.[4]

(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 not know I'm living in/speaking about a countable model?". The present family of views does not have this flavour, accepting instead that we can perfectly well refer to the universe, it is just that the level of **Forcing Absoluteness** is so strong that we can only talk about 'uncountable' sets by missing out functions.)

A different but closely related way to incorporate **Forcing Absoluteness** and keep the central role of ZFC is to deny **Capture**. One such position, proposed by [Steel, 2014] and taken up by [Meadows, 2015], doesn't allow for climbing extensions but does allow for forcing extensions. So it is forcing, but not climbing, multiversist. Steel, however, wants to have ZFC at every world, so can't endorse **Forcing Absoluteness** at that world in its full generality.

What one can do is *reinterpret* the **Forcing Absoluteness** to say that if we have a partial order in a universe, then there is *some* world in which there is a generic for that partial order. Let's just talk a little more about the details of Steel's theory to get a picture of this. Steel has a two-sorted theory with variables for sets $x_0, x_1, ...$ and variables for universes $W_0, W_1, ...$ with the following axioms (here I follow the presentation in [Maddy and Meadows, 2020]):

Definition 38. *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.

[4]See [Barton and Friedman, U] for these results (for the cognoscenti: One can get $0^\sharp$. The fact that we get large cardinals in inner models also results in some pleasant maximisation properties, see [Barton, Sb] for details.

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

So recalling our original statement of **Forcing Absoluteness**:

Forcing absoluteness. If there is a forcing extension such that $\exists x\phi$, then $\exists y\phi$.

Then under Steel's view, if there is a forcing extension with a set x such that ϕ, we can make a y such that $\phi(y)$ simply by moving to that extension and letting y be x. So instead of denying Powerset, we deny **Capture** and reinterpret **Forcing Absoluteness**.

One might think then, that actually the resemblance between Steel's approach and the view proposed in [Barton and Friedman, U] are rather different. Here it's important to recall the Mirroring Theorems. In particular since Steel's theory satisfies S4.2, there is a mirroring theorem available. In particular the translations into a non-modal theory will include both the claim that every set is countable (since for any set x it is possible that x is countable) and that there are many inner models for ZFC plus large cardinals (as long as we assert that there are worlds satisfying these axioms).[5]

Before we contrast this with the non-modal theory given in [Barton and Friedman, U], we should discuss another theory that yields similar results. [Scambler, 2021] has proposed a modal theory that is forcing multiversist but *also* climbing multiversist. For Scambler, not only can we always force, but given any conception of the second-order variables over some world, we can also reify all of them into a set. (Exactly how this is formalised is a little fiddly and requires meticulous attention to detail, so I will just direct the reader to [Scambler, 2021] for the specifcs.) Scambler then shows that under the potentialist translation, we have again ZFC minus Powerset with every set countable, again because any set found in some world can be collapsed.

It's now useful to contrast the case of these programmes with the ones given in the previous section. There we saw that both the Universist and Climbing Multiversist pictures converged towards a common picture—one that is strongly iterative and has as the base theory ZFC plus large cardinals, with a restricted amount of **Forcing Absoluteness** thrown in if desired. Here we have a similar situation, all of [Steel, 2014] [Meadows, 2015], [Scambler, 2021], and [Barton and Friedman, U] converge towards a similar non-modal picture, that of ZFC minus Powerset with every set countable, and ZFC plus large cardinals holding in

[5] I cover this in more detail in [NAB: Manuscript to be added.].

inner models (if desired). The conceptual picture thus advanced is similar, even if the details seem quite different.

Let's briefly consider how the «iterative conception» behaves on this picture, and what kind of **Theory of Infinity** is proposed. The first thing to note is that [Steel, 2014]/[Meadows, 2015] and [Scambler, 2021] are *weakly* iterative. On each there is an operation that is iterated to obtain sets. For Steel it is starting with a world and progressively forcing. This has the interesting consequence that the 'stages' provided, and from which we form new sets, are proper-class-sized. For Scambler it is starting with a world and adding ranks and forcing. So on the Scambler view, every stage forms a set in some other world. Every set, according to these pictures, is obtained this way, perhaps starting with different base worlds (it is unclear that there is a canonical choice of base world as there is for the advocate of the «climbing absoluteness conception» , where the Powerset Axiom holds and we can always start with $\emptyset$). But we have a description of some starting objects and an account of how new sets are obtained from old via some kind of process. So both the advocate of the «climbing absoluteness conception» and «forcing absoluteness conception» have iterative accounts of how their sets are obtained.

(**Note:** This analogy isn't quite perfect, since on the strongly iterative picture we also have the theorem in ZFC that every x is found in some V_α. It's not quite clear what such a theorem would look like on the weakly iterative picture.)

However, the «forcing absoluteness conception» is *not* strongly iterative, on pain of contradiction. Neither picture can obtain *all possible* subsets at an additional stage, since (in particular) there's always new subsets of the natural numbers (i.e. reals) that are outside a given stage. So the «forcing absoluteness conception» is very different from the «climbing absoluteness conception» in this regard. The only sense to be given to *strong* iterativity is from *within* a world/stage that can satisfy ZFC (and hence has its own conception of the V_αs).

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

Theory of Infinity for ZFC**.** How should we understand the **Theory of Infinity** provided by ZFC.

'Real' Theory of Infinity. What is the *real* **Theory of Infinity**?

The **Theory of Infinity for** ZFC is easily handled under the «forcing absoluteness conception» —the behaviour of the continuum function should be understood via the diverse world-to-world information we get out of forcing. As Joel-David Hamkins puts it:

> On the multiverse view, consequently, 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, for the «forcing absoluteness conception» , there is no maximal ZFC world, we have an answer to the question of how sets behave in ZFC, the answer is to be found in how it behaves across the worlds that satisfy ZFC. No further answer is needed or possible.

There is, however, the question of what the **'Real' Theory of Infinity** is. This question is answered for *sets*—every set is either finite or countably infinite. Since the continuum is a proper class (in fact, since you can think of any countable set as coded by a real, you can think of the universe as coded *by* the continuum and vice versa). 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 **Forcing Absolutist** must address.

Note: 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.

9.6 The schematic conception

Before moving on, I want to consider the «schematic conception» of MAXIMAL ITERATIVE SET. This simply says that *any* first-order model of ZFC is a legitimate universe of set theory, ontologically on a par with any other. It is normally also supplemented with climbing and forcing multiversism, yielding many models of set theory.

Schematic multiversism however goes much further. Since there are non-standard models of ZFC, it also allows for many non-well-founded universes of set theory as legitimate. These models still have a conception of the V_α and believe that every set belongs to some V_α. They thus still have the *mathematical* content of the iterative conception encoded within each universe. *Philosophically* however, there is no account of how sets are successively collected to form new ones. The account is rather *algebraic*, asserting that the subject matter of set theory is simply whatever structures satisfy some axioms (i.e. ZFC) and how these structures relate to one another, much as group theory is the study of structures satisfying the group axioms.[6]

Perhaps this is a philosophical problem that can be remedied, and there is some way of thinking of the «schematic conception» via methods of successive individuation. Since I do not see an obvious way out, I'll set this aside—the approach seems a radical departure from either the «climbing absoluteness conception» or «forcing absoluteness conception» .

9.7 Inference to the best conception and contemporary engineering

Where does all this leave us with respect to inference to the best conception? Both the «climbing absoluteness conception» and «forcing absoluteness conception» of MAXIMAL ITERATIVE SET have different ways of responding to **Theory of Infinity** and advocate very different responses. Both have some open questions to answer, but resolve many independent questions regarding large cardinals and the size of the continuum.

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 and the method of inference to the best conception. I won't come down one way or the other here—I think there are many questions to be left open for the future. However, it's good to examine some features of the two views moving forward. Recall, the main foundational goals we looked at (aside from **Theory of Infinity**, which we've already discussed) were:

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

[6] I examine this conception in more detail in [Barton, 2016].

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.

Generous arena is handled very differently by the two approaches. But each has their own answer. The «climbing absoluteness 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 absoluteness conception» is a little more interesting. Here 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* within the non-modal theory. Whether this constitutes a hobbling of mathematical practice or just a different approach is a question I leave open for philosophical examination (I provide a little more mathematical detail in [Barton and Friedman, U]).

This has implications for **Shared Standard**. Both the «forcing absoluteness conception» and «climbing absoluteness conception» provide their own **Generous Arena**, and hence their own account of when a proof is legitimate. Each is very different though, the «forcing absoluteness conception» says that third-order resources are not legitimate for reasoning about the reals. So both have an account of **Shared Standard**, but the «forcing absoluteness conception» deviates substantially from the currently accepted norm. This said, under the «forcing absoluteness conception» proofs in third-order arithmetic and/or ZFC are not *wrong*, they just need to be interpreted in *restricted contexts*. Again, whether this should count *against* the «forcing absoluteness conception» or it is simply merely *different*, I leave open.

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[7]). 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 «climbing absoluteness conception» provides an answer to many questions about large cardinal independence, and (when combined with a consistent amount of forcing absoluteness) resolves CH but leaves many other questions unanswered. The «forcing absoluteness 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 $F : V \rightarrowtail\!\!\!\!\rightarrow Ord$) . So whilst the sets are relatively easily known, the continuum is not, and is in fact connected with the nature of proper classes.

Moreover, both provide interesting perspectives as a **Testing Ground for Paradox**. This is in two ways. First, the incompatability between **Climbing Absoluteness** and **Forcing Absoluteness** 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). Moreover, although each denies the full generality of the other's principles, one can incorporate *partial amounts* thereof. The proponent of the «climbing absoluteness conception» can add in limited amounts of **Forcing Absoluteness** (indeed this is needed for the resolution of CH via bounded forcing axioms), and vice versa (the advocate of the «forcing absoluteness conception» will likely want to add in *some* climbing absoluteness, even if the amount allowed is rather limited[8]).

However, each provides their own problems of extension to inconsistency. We haven't really covered this in detail here, since our focus has been on drawing out the incompatibility between the two views. However it bears mentioning that each kind of absoluteness can be extended to inconsistency. For instance, **Climbing Absoluteness** is inconsistent if we allow reflection with unrestricted third-order parameters.[9] **Forcing Absoluteness** is inconsistent when extended to allow for wider kinds of extensions (in particular class forcing) and unre-

[7]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.

[8]Specifically, Dependent Choice can be formulated as a climbing absoluteness principle. See [Barton and Friedman, U] for discussion.

[9]See here [Reinhardt, 1974], [Tait, 2005], and [Koellner, 2009].

stricted parameters.[10] So each view is philosophically interesting—presenting their own kinds of 'paradox' via extension to inconsistency.

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 large cardinal 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 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 MAXIMAL ITERATIVE SET.

For these reasons I think that both the «forcing absoluteness conception» and «climbing absoluteness conception» are viable conceptions of set. The «climbing absoluteness conception» clearly fits better with current orthodoxy, but that's not a good reason to discount the «forcing absoluteness conception» out of hand. In the end, I think that a careful analysis via the method of inference to the best conception 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 «forcing absoluteness conception» . Before we wrap up in Chapter 11, I want to consider some objections.

[10] See [Barton and Friedman, U], Theorem 25.

Chapter 10

Objections

This chapter has one very simple aim: Examine and respond to what I take to be the most salient objections to what I've proposed here.

10.1 No one has ever held the absoluteness conception

I first want to consider an objection raised earlier, and reiterate why I think it's not a problem. The «absoluteness conception» , one might argue, is just *obviously* inconsistent, and no-one has ever seriously advocated it. Set theorists are a bright bunch of folks, and they can see the contradiction of the Cohen-Scott paradox coming before we even get started.

It is correct that no-one single agent has advanced the absoluteness conception in the literature (though I presume some people have fallen into the trap when learning forcing). What I want to contend is that the community is being pulled in two orthogonal directions by a desire to have 'as many sets as possible', even if no-one individual agent is advocating the conception.

It's perhaps useful to contrast this situation with that of our ancestors in the period from the discovery of the paradoxes until about 1950. There, they were clearly aware that we couldn't have the «naive conception» . No one agent was advocated both an **Indefinite Extensibility** and **Universality** approach after the discovery of the paradoxes (I assume no-one was a dialethist until relatively recently). However, as discussed earlier (Chapter 5) aspects of the «iterative conception» and «stratified conception» were latent in their thought. But they lacked the conceptual resources to know what precise conceptions would emerge, and lacked the tools to express themselves clearly enough.

Clearly our tools have come a long way, and the resources of formal logic allow us to express ourselves clearly and identify contradictions and different conceptions with more precision. But we (as a community) are still pulled in different directions concerning our thoughts about absoluteness, and it may be that a new perspective is needed in order to see our intellectual predicament more clearly.

10.2 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 they are either correct or incorrect about this universe, and this is the only arbiter of truth we need.

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? Do we really think, out of all the possible conceptions we might have, that we've selected the 'right' one? I should add, there are many more conceptions out there (see [Incurvati, 2020]) and these are probably just 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.[1]

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. Better, I think, to discuss what we want to *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.

10.3 The iterative conception is just fine, thanks

A different response is to argue that the whole project with which I'm engaged is uninteresting because ZFC and the «iterative conception» work just fine. The point might be developed by pointing to

[1] I also make this argument in [Barton, 2022].

the fact, outlined in Chapter 2, that ZFC suffices for almost all known mathematics.

I think this response is fine, as far as it goes. If all you want out of a foundation is to provide you with the bare bones of what's required for most of concrete mathematics, ZFC more than suffices. If so motivated, one might even stop at or before $V_{\omega+\omega}$. But we are after a good and informative foundation that does more, and in particular if you a decent **Theory of Infinity** you're going to have to engage with these difficult questions.

An analogy here is useful here between the case of MAXIMAL ITERATIVE SET and TRUTH as expounded by [Scharp, 2013]. Scharp argues that for most applications, TRUTH is fine, and it's only for annoying cases that we get drawn into paradox. We can point to other similar cases, MASS is a fine example. Standardly conceived, MASS is an inconsistent concept (at this world) though RELATIVISTIC MASS and PROPER MASS are not. [NAB: more here?] For most applications (e.g. the vast majority of engineering) MASS does just fine. But if you want to calibrate atomic clocks on satellites you'll need to take into account the difference. Similar remarks can be made about the «naive conception» of SET—for most applications its fine, but if you want to start seriously engaging with the foundations of set theory, one needs to pay attention to the distinction between the «iterative conception» and «stratified conception» (among others). I think the same is true here. For most applications the «iterative conception» is fine, but if you want to get into the finer details of providing a **Theory of Infinity** and what kind of theory is suggested by contemporary set-theoretic practice, it's good to think in terms of the «forcing absoluteness conception» and «climbing absoluteness conception» .

10.4 The story is too neat, and ignores much

This idea (that we should think of contemporary set-theoretic practice in terms of the «forcing absoluteness conception» and «climbing absoluteness conception») can be challenged. what about, for instance, inner model theory and the Ultimate-L programme? 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. 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 the «climbing absoluteness conception» and «forcing absoluteness conception» as *the* two possibilities for set-theoretic development, but rather simply indicate what kinds of choices there may be and how they are weighed. My focus was articulating a position on which we could see at least the two paths outlined as ways of proceeding, and highlight some similarities our own predicament and that of our intellectual ancestors. In particular, I made simplifying assumptions there too—there's far more than the «iterative conception» and «stratified conception», [Incurvati, 2020] highlights also the «graph conception», «limitation of size conception», and «naive dialethic conception» too. So the situation is far more complex than I let on, and a full examination will require a massive effort from historians, philosophers, and sociologists of mathematics, as well as philosophically interested mathematicians.

I do want to make one point about the Ultimate-L programme for the cognoscenti though. It is unclear to me whether this should be viewed as a conception of MAXIMAL ITERATIVE SET or just ITERATIVE SET simpliciter. We might view the programme as a suggestion that we do away with maximality considerations and instead focus on what they regard as more mathematically tractable ideas. If this is correct, then it indicates an interesting (different) kind of engineering route!

10.5 The 'change of subject' objection

A next criticism is one common to the conceptual engineering literature, and concerns whether or not we simply 'changed the subject' by engineering set-theoretic concepts/conceptions. Often traced back to Strawson's criticism of Carnapian explication (cf. [Strawson, 1963]), the objection in the current context would run roughly as follows: I've argued that we should understand us as engineering different set-theoretic conceptions by carefully calibrating the constitutive principles we wish to add. But one might think that any difference if constitutive principles yields a change in subject matter (after all, the intension of the concept/conception has changed). But—so the objection goes—perhaps this vitiates the possibility of seeing the present study as a genuine inquiry into questions like CH, instead there are different versions of CH for each conception, since the subject matter is different.

I want to make a few observations about how forceful one finds this argument. The first is to note that this problem is not unique to set theory, but is a general complaint about conceptual engineering. I'm don't have space to discuss the problem in detail, but we can say a few things about the specific case of set theory.

The first is that many of the standard responses to this problem will transfer to the current setting. One, proposed by [Cappelen, 2018], suggests that we take subject matter to be individuated by 'topics'. Two utterances u_0 and u_1 have the same topic if they can be can be embedded in other utterances giving the a cross contextual report (a relation Cappelen calls 'samesaying'). So in a present context, let's take an agent A_F operating with the «forcing absoluteness conception» and an agent A_C operating with the «climbing absoluteness conception», and suppose that each utters ¬CH. Is the utterance "A_F and A_C both said that CH is false" felicitous?

I don't see any good way to get traction on this issue beyond the intuition of whether there is continuity of content simpliciter. It seems at least *plausible* that such reports are possible. This is perhaps clearer if we look at agents involved in engineering in the past. To my mind, the statement "Both Cantor and Maddy are interested in the status of the Continuum Hypothesis" seems like a perfectly legitimate claim to make, even though engineering has occurred between Cantor's time and Maddy's (and, as I've argued, they're not working with the same conception of SET).

NAB: The above paragraph is a little unsatisfying, but it's where I'm at.

A different response is to argue that there's a continuity of subject matter when there's a continuity of *functional role*. This approach is prevalent in the literature on social justice. Sally Haslanger, for example, argues that whilst FAMILY has undergone substantial engineering, there is a continuity of subject matter in virtue of the fact that FAMILY fills largely the same functional role (by giving us information about the coordination of domestic life) even though the intension and extension of the concept has changed. If we look at the case of SET (and the «forcing absoluteness conception» and «climbing absoluteness conception») we see that (by the discussion of the previous chapter, especially §9.7) both aim at fulfilling similar foundational roles, even if they do so in competing ways. In fact, the standards outlined in Chapter 2 entail that any conception meeting those standards is likely to fulfil a similar functional role.

The question of whether one finds the 'change of subject matter' objection convincing can be articulated by what Delia Belleri calls semantically *conservative* versus semantically *progressive* inquiries (cf. [Belleri, 2021]). The former demand semantic continuity, whereas the latter allow for semantic change. Belleri asks us to consider the question of whether fish have gills (as asked by a curious child), and contrast it with the question of whether Pluto is a planet. The plutonic question (in contrast to the fishy one) admits of semantic change, since there are various pragmatic questions surrounding whether to include Pluto as a planet (such as then being forced to rule in several more large-ish objects in the Kuiper belt as planets, making our demarcation unwieldy). The question is whether set theory is closer to the latter or former kind of inquiry. Given that engineering has already happened, it seems to me that set theory is a semantically progressive discipline, or at least the claim that it is not has to be justified, and the case for why our situation is not like that of our intellectual forebears made out.

10.6 How are set theorists able to agree?

One way the previous argument might be strengthened is by showing that there is communicative discontinuity. If there's really multiple different conceptions of set in play, and different parts of the community are operating with these different conceptions, then how are set-theorists able to agree? After all, set-theorists from a wide variety of conceptual schools are able to evaluate and understand each others work. But if they mean different things with their use of the term "set" then how does communication not break down?

This is an interesting question, and a full answer will have to wait. Really we should get into the finer points of the nature of concepts and conceptual agreement, and there's just not space to do that here. However, let me indicate a way this worry can be assuaged. The core point is that both the advocate of the «forcing absoluteness conception» and «climbing absoluteness conception» can see the other as talking about substructures of their world, and indeed interesting substructures thereof. The advocate of the «climbing absoluteness conception» sees the advocate of the «forcing absoluteness conception» as essentially studying analysis/second-order arithmetic using set-theoretic tools—an interesting project from their perspective! And the advocate of the «forcing absoluteness conception» sees the advocate of the «climbing absoluteness conception» as studying inner models or countable transitive models of the universe, which they also regard as very important for their project (especially if they ascribe to a modal weakly iterative story). Both are able to examine the correctness of proofs provided by their interlocutors in these terms. It is just when it comes to the final judgement about what is *true*, they may differ.

Moreover, especially mobile theorists will be able to move from one conception to another. Just as we remarked (way back in Chapter 3) that Jane possesses the concept of FAIRNESS-BY-OUTCOME even though she holds the «fairness-by-effort conception» of FAIRNESS, and thereby can reason about what holds for the concept FAIRNESS-BY-OUTCOME, so can an advocate of one of our two set-theoretic conceptions possess the concept of the other and reason about what happens under that concept. Possessing a concept is a different thing from holding that it is the right conception, and being able to reason in these terms is essential for being part of a thriving intellectual community.

Chapter 11

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. Moreover, I think that by examining the history of set theory, we can see instances of conceptual engineering where there is a trade-off between inconsistent constitutive principles, and different conceptions can be isolated by choosing one over the other. Moreover, 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. So, to start:

Can Capture be a constitutive principle? Throughout, I've acted as though **Capture** is not a constitutive principle, but rather a metaphysical claim that can figure into how claims are cashed out in various ways. However, when we examine what I've argued—for example how the «forcing absoluteness conception» is likely to result in different (but intertranslatable) theories depending on ones attitude to **Capture**—we can see that different *languages* and *theories* are naturally supported by different attitudes to **Capture**. This suggests that we might view **Capture** as a constitutive principle in its own right, yielding different conceptions (e.g. the «**Capture**-validating forcing absoluteness conception » of MAXIMAL ITERATIVE SET etc.). We should be careful though—presumably we don't want any hodgepodge collection of beliefs to count as a conception. So more attention is needed regarding the status of **Capture**, and with it how fine-grained conceptions can be.[1]

What about other conceptions? A salient point concerning the present study is that we went very narrow quite quickly, moving directly to the «iterative conception» and then the «absoluteness conception» and its sub-conceptions. There's several ways that the present way of thinking about conceptions and conceptual engineering could be broadened. For example, we didn't consider other, non-iterative, conceptions of SET. So we might ask: **What of other conceptions of SET (e.g. those considered by [Incurvati, 2020])?** What might also ask how broad this can be, asking questions like: **Are there conceptions of SET that validate logics weaker than classical logic?**[2] Moreover, a central question that I didn't examine much was whether there are different 'rival' conceptions *within* ZFC (given that the «forcing absoluteness conception» pushes us toward every set being countable). We can ask: **Are there multiple competing conceptions of ITERATIVE SET when we fix ZFC?** (A good example here that merits serious further consideration is the Ultimate-L programme.) And how does all this figure into the present mathematical landscape?

What of open texture?[3] *Open texture*, as it appears in [Waismann, 1947], is the idea that our concepts might not effectively categorise objects, and in fact it might be that objects can be placed under different concepts according to contingent facts. An example from [Wilson, 2006]: A group of people who have never seen a plane before might classify a plane as a kind of strange bird or a kind of house, depending on whether they first see the plane flying in the sky, or the crew using the downed wreckage as a place of abode. In this way, the concepts HOUSE and BIRD exhibit *open texture*, the plane might or might not fall under them, depending on other factors about how the object is presented. Originally, this kind of phenomenon was thought to not be possible for mathematical concepts. However, recently Stewart Shapiro has argued that computability provides a mathematical example (cf. [Shapiro, 2013]). In particular he argues that the original notion of 'effective procedure' exhibited open texture, before finally being solidified by the Church-Turing Thesis. So it's at least possible that we have a case of open texture regarding SET. But do we?

[1] A talk version of some of this material given in Oslo originally had **Capture** as a constitutive principle, and I'm grateful to Herman Cappelen and Øystein Linnebo for some discussion here. I am at least a little sympathetic to the idea that conceptions might be very fine-grained, but that these very fine-grained conceptions form natural 'clusters'. This goes strongly against [NAB: Cite Quine.].

[2] One suggestion here, considered by [Incurvati, 2020], is [Priest, 1995]'s paraconsistent theories concerning the «naive conception» .

[3] I'm grateful to Stewart Shapiro for some discussion of this question.

I think that the situation is complicated, and made additionally more complex by the fact that when we consider the concept EFFECTIVELY COMPUTABLE FUNCTION, we are looking to apply this notion an *already given* domain of natural numbers and functions, and separate some functions out as the effectively computable ones. By contrast, when we look at conceptions of SET, we are in part *deciding what to talk about*. Of course, one can *enforce* open texture, say by adopting some kind of structuralism and holding that we are classifying different *structures* as falling under or outside the concept SET. Still it's very unclear that we can straightforwardly apply the Shapiro-model to the case of SET.

What of the radical therapeutic programme? In [Scharp, 2013] and [Scharp, 2020], Kevin Scharp suggests that we should engage in a kind of 'radical therapeutic programme', where the job of the philosopher is to isolate different scientifically tractable concepts for study. These concepts can then be given to the scientist, and the philosopher need not concern themselves with the concept any further.

I want to suggest that what we have here might fit this mould to some extent, but in a more symbiotic fashion. To my mind, concepts can be developed by the philosopher and then handed over to the scientist, a process that naturally suggests more philosophical concerns for consideration. This process is suggested by the present study, we can view both the «climbing absoluteness conception» and «forcing absoluteness conception» (as well as other conceptions) as now up for mathematical consideration. No doubt this will suggest further refinements. But aside from this rough sketch, there is much further work to be done in this direction.

What of pluralism? I want to close with the following question: We've seen that historically there was engineering, and inference to the best conception combined with social factors resulting in the «iterative conception» coming out on top (though perhaps other conceptions are experiencing something of a resurgence). Moreover, I've argued that we now find ourselves at a fundamental choice point, do we go with the «forcing absoluteness conception», the «climbing absoluteness conception», or something else entirely? However, there's a third option: We might end up in a situation in which the different conceptions perform better with respect to certain criteria compared to others. This would result in 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 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 for the «forcing absoluteness conception» . To me, it seems philosophically open whether or not the view of the continuum as something much 'larger' than was previously envisaged is really conceptually worse than the usual ZFC-based rival. We can only hope to solve these issues by thoughtful mathematical and philosophical analysis of different conceptions, contrasting their features and analysing the agreement between them. Even then, it's not clear how much control we have over our semantic whims.[4] The future is open and exciting, with a good deal of work to be done in understanding our place with respect to the world(s) of infinite sets.

[4]The idea that we don't have much control is advocated by [Wilson, 2006].

Bibliography

[Antos, 2015] Antos, C. (2015). *Foundations of Higher-Order Forcing*. PhD thesis, Universität Wien.

[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.

[Antos and Friedman, 2017] Antos, C. and Friedman, S.-D. (2017). Hyperclass forcing in Morse-Kelley class theory. *The Journal of Symbolic Logic*, 82(2):549–575.

[Arrigoni and Friedman, 2013] Arrigoni, T. and Friedman, S.-D. (2013). The Hyperuniverse Program. *Bulletin of Symbolic Logic*, 19:77–96.

[Bagaria, 2000] Bagaria, J. (2000). Bounded forcing axioms as principles of generic absoluteness. *Archive for Mathematical Logic*, 39(6):393–401.

[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, 2016] Barton, N. (2016). Richness and reflection. *Philosophia Mathematica*, 24(3):330–359.

[Barton, 2020] Barton, N. (2020). Forcing and the universe of sets: Must we lose insight? *Journal of Philosophical Logic*, 49(4):575–612.

[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, Sa] Barton, N. (Sa). Are large cardinal axioms restrictive? Manuscript under review. Pre-Print: `https://philpapers.org/archive/BARALC-5.pdf`.

[Barton, Sb] Barton, N. (Sb). Is (un)countabilism restrictive? Manuscript under review. Pre-Print: FIXIFIX.

[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, U] Barton, N. and Friedman, S. (U). Countabilism and maximality principles. Manuscript under review.

[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.

[Barton and Williams, U] Barton, N. and Williams, K. (U). Varieties of class-theoretic potentialism. Manuscript under review. Preprint: `https://philpapers.org/rec/BARVOC-4`.

[Barwise, 1975] Barwise, J. (1975). *Admissible Sets and Structures: An approach to definability theory*. Springer.

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

[Belleri, 2021] Belleri, D. (2021). Downplaying the change of subject objection to conceptual engineering. *Inquiry*, 0(0):1–24.

[Benacerraf and Putnam, 1984] Benacerraf, P. and Putnam, H., editors (1984). *Philosophy of Mathematics: Selected Readings*. Cambridge University Press.

[Boolos, 1971] Boolos, G. (1971). The Iterative Conception of Set. *The Journal of Philosophy*, 68(8):215–231.

[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*.

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

[Button and Walsh, 2018] Button, T. and Walsh, S. (2018). *Philosophy and Model Theory*. Oxford University Press.

[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.

[Cantor, 1883] Cantor, G. (1883). Foundations of a general theory of manifolds; a mathematico-philosophical investigation into the theory of the infinite. In Ewald, W., editor, *From Kant to Hilbert; A Source Book in the Foundations of Mathematics*, volume 2. Oxford University Press.

[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.

[Cocchiarella, 1985] Cocchiarella, N. B. (1985). Frege's double correlation thesis and quine's set theories nf and ml. *Journal of Philosophical Logic*, 14(1):1–39.

[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.

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

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

[Ewald, 1996] Ewald, W. B., editor (1996). *From Kant to Hilbert. A Source Book in the Foundations of Mathematics*, volume I. Oxford University Press.

[Field, 1980] Field, H. H. (1980). *Science Without Numbers: A Defence of Nominalism*. Princeton, NJ, USA: Princeton University Press.

[Field, 1989] Field, H. H. (1989). *Realism, Mathematics & Modality*. Blackwell.

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

[Friedman, 2006] Friedman, S.-D. (2006). Internal consistency and the inner model hypothesis. *Bulletin of Symbolic Logic*, 12(4):591–600.

[Friedman et al., 2008] Friedman, S.-D., Welch, P., and Woodin, W. H. (2008). On the consistency strength of the inner model hypothesis. *The Journal of Symbolic Logic*, 73(2):391–400.

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

[Gitman and Schindler, 2018] Gitman, V. and Schindler, R. (2018). Virtual large cardinals. *Annals of Pure and Applied Logic*.

[Gödel, 1964] Gödel, K. (1964). What is Cantor's continuum problem? In *[Gödel, 1990]*, pages 254–270. Oxford University Press.

[Gödel, 1990] Gödel, K. (1990). *Collected Works, Volume II: Publications 1938-1974*. Oxford University Press. Edited by: Solomon Feferman (Editor-in-chief), John W. Dawson, Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, Jean van Heijenoort.

[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 Linnebo, 2022] Hamkins, J. D. and Linnebo, O. (2022). The modal logic of set-theoretic potentialism and the potentialist maximality principles. *The Review of Symbolic Logic*, 15(1):1–35.

[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.

[Holmes, 1998] Holmes, R. (1998). *Elementary Set Theory with a Universal Set*. Bruylant-Academia.

[Hossack, 2014] Hossack, K. (2014). Sets and plural comprehension. *Journal of Philosophical Logic*, 43(2-3):517–539.

[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.

[Jech, 2002] Jech, T. (2002). *Set Theory*. Springer.

[Kanamori, 1996] Kanamori, A. (1996). The mathematical development of set theory from Cantor to Cohen. *The Bulletin of Symbolic Logic*, 2(1):1–71.

[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.

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

[Leitgeb and Carus, 2021] Leitgeb, H. and Carus, A. (2021). Rudolf Carnap. In Zalta, E. N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2021 edition.

[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.

[Maddy, 1998] Maddy, P. (1998). $V = L$ and maximize. In Makowskyf, J. A. and Ravve, E. V., editors, *Proceedings of the Annual European Summer Meeting of the Association of Symbolic Logic*, pages 134–152. Springer.

[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.

[Magidor, 2012] Magidor, M. (2012). Some set theories are more equal. In Koellner, P., editor, *Exploring the Frontiers of Incompleteness*. Preprint available from `http://logic.harvard.edu/EFI_Magidor.pdf`.

[Malliaris and Shelah, 2016] Malliaris, M. and Shelah, S. (2016). Cofinality spectrum theorems in model theory, set theory, and general topology. *Journal of the American Mathematical Society*, 29:237–297.

[Martin, 1970] Martin, D. (1970). Review of *Set Theory and its Logic* by Willard Van Orman Quine. *Journal of Philosophy*, 67:111–114.

[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.

[Mirimanoff, 1917] Mirimanoff, D. (1917). Les antinomies de Russell et de Burali-Forti et le probleme fondamental de la theorie des ensembles. *L'Enseignement Mathématique*, 19:37–52.

[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, 1995] Priest, G. (1995). *Beyond the Limits of Thought*. Oxford University Press.

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

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

[Reinhardt, 1980] Reinhardt, W. (1980). Satisfaction definitions and axioms of infinity in a theory of properties with necessisty operator. In Arruda, A., Chuaqui, R., and Costa, N. D., editors, *Mathematical Logic in Latin America*, pages 267–303. North Holland Publishing Co.

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

[Roberts, 2016] Roberts, S. (2016). *Reflection and Potentialism*. PhD thesis, Birkbeck, University of London.

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

[Rosser, 1953] Rosser, J. B. (1953). *Logic for Mathematicians*. McGraw-Hill.

[Russell, 1959] Russell, B. (1959). *My Philosophical Development*. Routledge.

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

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

[Scharp, 2020] Scharp, K. (2020). Philosophy as the study of defective concepts. In *Conceptual Engineering and Conceptual Ethics*, pages 396–416. 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, 2013] Shapiro, S. (2013). The open texture of computability. In Copeland, B. J., Posy, C. J., and Shagrir, O., editors, *Computability: Turing, Gödel, Church, and Beyond*, pages 153–181. MIT Press, Cambridge, MA.

[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.

[Shoenfield, 1967] Shoenfield, J. (1967). *Mathematical Logic*. Addison-Wesley Publishing Co.

[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.

[Soysal, 2020] Soysal, Z. (2020). Why is the universe of sets not a set? *Synthese*, 197(2):575–597.

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

[Strawson, 1963] Strawson, P. F. (1963). Carnap?s views on conceptual systems versus natural languages in analytic philosophy. In Schilpp, P. A., editor, *The Philosophy of Rudolf Carnap*, pages 503–518. Open Court: La Salle.

[Tait, 2005] Tait, W. (2005). Constructing cardinals from below. In *The Provenance of Pure Reason: Essays in the Philosophy of Mathematics and Its History, Logic and Computation in Philosophy*, pages 133–154. Oxford University Press.

[Tanswell, 2018] Tanswell, F. S. (2018). Conceptual engineering for mathematical concepts. *Inquiry*, 61(8):881–913.

[Turing, 1937] Turing, A. M. (1937). On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society*, s2-42(1):230–265.

[Uzquiano, 2003] Uzquiano, G. (2003). Plural quantification and classes. *Philosophia Mathematica*, 11(1):67–81.

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

[Waismann, 1947] Waismann, F. (1947). Verifiability. *Journal of Symbolic Logic*, 12(3):101–101.

[Wang, 1984] Wang, H. (1984). The concept of set. In *[Benacerraf and Putnam, 1984]*, pages 530–570. Cambridge University Press.

[Welch and Horsten, 2016] Welch, P. and Horsten, L. (2016). Reflecting on absolute infinity. *Journal of Philosophy*, 113(2):89–111.

[Williams, 2018] Williams, K. J. (2018). *The structure of models of second-order set theories*. PhD thesis, The Graduate Center, The City University of New York.

[Williams, 2019] Williams, K. J. (2019). Minimum models of second-order set theories. *The Journal of Symbolic Logic*, 84(2):589–620.

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

[Zermelo, 1930] Zermelo, E. (1930). On boundary numbers and domains of sets. In *[Ewald, 1996]*, volume 2, pages 1208–1233. Oxford University Press.