Forcing and the Universe of Sets: Must we lose insight?

Neil Barton*

7 June 2019†

Abstract

A central area of current philosophical debate in the foundations of mathematics concerns whether or not there is a single, maximal, universe of set theory. Universists maintain that there is such a universe, while Multiversists argue that there are many universes, no one of which is ontologically privileged. Often *forcing* constructions that add subsets to models are cited as evidence in favour of the latter. This paper informs this debate by analysing ways the Universist might interpret this discourse that seems to necessitate the addition of subsets to V. We argue that despite the prima facie incoherence of such talk for the Universist, she nonetheless has reason to try and provide interpretation of this discourse. We analyse extant interpretations of such talk, and analyse various tradeoffs in naturality that might be made. We conclude that the Universist has promising options for interpreting different forcing constructions.

Introduction

Recent discussions of the philosophy of set theory have often focussed on how many universes of sets there are. The following is a standard position:

> **Universism.** There is a unique, maximal, proper-class-sized universe containing all the sets (denoted by 'V').

Universism has often been thought of as the 'default' position on the ontology of sets.[1] However, some have seen the existence of many different epistemic possibilities for the nature of the set-theoretic universe, shown by the large diversity of model-theoretic constructions witnessing independence (we discuss two of these methods later) as indicative of the existence of a diversity of different set-theoretic universes. In this paper, we will be concerned with *forcing*. This technique has been undeniably central in the study of independence in set theory, and (prima facie at

*Zukunftskolleg, University of Konstanz. E-mail: neil.barton@uni-konstanz.de.

†The author wishes to thank Carolin Antos, Tim Button, Andrés Eduardo Caicedo, Monroe Eskew, Sy Friedman, Vera Flocke, Victoria Gitman, Bob Hale, Joel Hamkins, Toby Meadows, Sandra Müller, Alex Paseau, Ian Rumfitt, Chris Scambler, Sam Roberts, Jonathan Schilhan, Daniel Soukup, Kameryn Williams, and three anonymous reviewers for insightful and helpful comments, as well as audiences in Cambridge, Konstanz, Vienna, and Toulouse for the opportunity to present and subsequent discussion. He is also very grateful for the generous support of the UK Arts and Humanities Research Council (as a PhD student), FWF (Austrian Science Fund, Project P 28420), and VolkswagenStiftung (through the project *Forcing: Conceptual Change in the Foundations of Mathematics*).

[1] Both Gödel and Cantor, for example, are notable proponents.

least) adds a set (or—as we shall see—a class) G to a model $\mathfrak{M}$ to form a *forcing extension* $\mathfrak{M}[G]$.

In broad strokes, the problem arises when we note that set-theoretic practice is replete with uses of forcing where the symbol 'V' is used to denote the model over which we force (the so-called 'ground model'). Since, for non-trivial forcing, the relevant G added cannot be in the ground model, one might naively think that the relevant G lives outside V, and so the Universist position is immediately false.

As is well-known, however, there are various interpretative strategies available to the Universist when it comes to forcing. We discuss these in detail later, but for now we note that two popular options have been the use of countable transitive models and Boolean-valued models. In these cases, we re-interpret some or all of the terms in the forcing construction to yield an interpretation of the forcing talk that does not commit us to sets outside V. Opponents of Universism sometimes argue that these interpretations are problematic.[2] In this paper, we argue for the following claims:

(1.) There is some pressure arising from set-theoretic practice to provide an interpretation of forcing where the use of 'V' in a forcing construction is linked to *actual* truth in V.

(2.) Nonetheless, the Universist has several options available, and only *mild* compromises have to be made for some of these interpretations.

In this way, we aim to advance and clarify the dialectic between the Universist and her opponents. In particular we will argue that underpinning (1.) are additional philosophical commitments to naturality of interpretation, and that these are (2.) largely satisfied when considering some extant possibilities for interpreting forcing constructions.

Here's the plan: After these introductory remarks, we (§1) briefly outline the kinds of forcing construction we will consider. Next (§2) we will argue that despite the fact that the Universist has methods for interpreting the forcing required for independence results, there is nonetheless some pressure to interpret forcing in such a way that its use is linked to the use of the symbol 'V' in forcing constructions. Specifically, we will argue that the Universist is able (through the formulation of axioms and proving of theorems) to learn more about the structure of V by viewing V as situated in a multiverse framework. We'll argue then that if the Universist accepts certain constraints on the naturalness of interpretation of mathematical discourse, there is additional pressure on her interpretation of forcing. Finally (§3) we provide an analysis of extant interpretations of forcing found in the literature (the forcing relation, Boolean-valued models, the Boolean ultrapower, and countable transitive models). We argue that use of the forcing relation and Boolean-valued models largely violate the constraints of §2. We will argue that the situation for the Boolean ultrapower and countable transitive models is more subtle. While there are tradeoffs to be made for each interpretation, for the countable transitive model strategy these are relatively minimal (especially given the existence of certain large cardinals), and for a certain class of forcing constructions the Boolean ultrapower is especially natural (and especially unnatural for a different class). We conclude with some open questions concerning the directions of debate for both the Universist and her critics.

[2] A good example here is [Hamkins, 2012].

1 Varieties of forcing

We start with a brief description of forcing constructions to make plain some of their mathematical properties that will be philosophically relevant later. Forcing comes in two main kinds; *set* forcing and *class* forcing.[3] We briefly sketch the constructions (deferring to the relevant technical literature where appropriate) noting their similarities and differences. The details will be familiar to specialists, but a feel for some of the properties of the constructions is necessary for seeing the challenges faced.

1.1 Set forcing

For *set forcing* we begin with a partial order with domain P, ordering $\leqslant_{\mathbb{P}}$, and maximal element $1_{\mathbb{P}}$, denoted by '$\mathbb{P} = \langle P, \leqslant_{\mathbb{P}}, 1_{\mathbb{P}}\rangle$', and have $\mathbb{P} \in \mathfrak{M}$ for some **ZFC** model[4] $\mathfrak{M}$. The relevant $p \in P$ are known as *conditions* and effectively operate by providing partial information about membership of the new object to be defined. We then, via a careful choice of names (known as '$\mathbb{P}$-*names*')[5], and evaluation procedure[6] add a filter G on $\mathbb{P}$ that intersects all dense sets of $\mathbb{P}$ in $\mathfrak{M}$ to $\mathfrak{M}$. The end result is a model $\mathfrak{M}[G]$ that (i) satisfies **ZFC**, (ii) has exactly the same ordinal height as $\mathfrak{M}$, and (iii) is strictly larger than $\mathfrak{M}$ (in the sense that $\mathfrak{M} \subset \mathfrak{M}[G]$).[7]

Forcing is an especially interesting philosophical construction for a number of reasons. First, it is historically significant in that it has been used to settle many open questions (the most famous examples being the independence of CH and AC). Second, it is of central importance in virtue of its ubiquity across modern set-theoretic mathematics; much of set theory concerns constructing one model from another using forcing arguments. However, especially *philosophically* interesting is that (for a wide class of structures) it keeps models *standard*[8]. Assuming that the ground model $\mathfrak{M}$ is transitive, well-founded, and satisfies **ZFC**, and that there is a generic G available, the forcing extension $\mathfrak{M}[G]$ (i) has the same ordinals as $\mathfrak{M}$, (ii) satisfies **ZFC**, and (iii) is transitive and well-founded. The fact that forcing often keeps the models *standard* is significant; generic extensions of a standard model of **ZFC** are also **ZFC**-satisfying cumulative hierarchies and hence 'look' like a bona fide universe of sets.

The issue concerning forcing and V is, of course, that if we wish to perform a non-trivial forcing where 'V' denotes the Universist's V as the ground model, the relevant generic G must lie outside V.[9] But V was meant to be all the sets there are,

[3] There is also the related notion of *arithmetic forcing* used to study models of second-order arithmetic. Since we are interested primarily in models of set theory here, we set this aside.

[4] A brief remark is in order here: One does not always force over models of full **ZFC**, and forcing over models of weaker theories is well studied. Indeed, in several mathematics texts (such as [Kunen, 2013]), the 'official' approach is to use the reflection theorem to obtain a model of 'enough' **ZFC** to conduct the independence proof. We defer consideration of this issue until later.

[5] 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 is not vacuous in virtue of the empty set trivially being a $\mathbb{P}$-name).

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

[7] It should be noted that in order for the forcing to be non-trivial, $\mathbb{P}$ has to be *non-atomic* (i.e. every $p \in P$ has incompatible extensions in $\mathbb{P}$).

[8] A model $\mathfrak{M}$ is normally called *standard* iff it has the real $\in$-relation. When there is a generic available, the resulting forcing extendion $\mathfrak{M}[G]$ will be standard if $\mathfrak{M}$ is. See [Kunen, 2013], §IV.2 for verification of the basic properties of forcing.

[9] This is because if G were in V, $\mathbb{P} - G$ would be a dense set in V missed by G.

and so such a G does not exist.

1.2 Class forcing

Class forcing is very similar to set forcing, except we drop the requirement that the partial order and generics are members of $\mathfrak{M}$ and instead permit *proper-class-sized* partial orders. The technique also uses partial orders with maximal elements $\langle \mathbb{P}, \leqslant_{\mathbb{P}}, 1_{\mathbb{P}} \rangle$, and adds a generic G to our ground model $\mathfrak{M}$. The difference with class forcing is that $\mathbb{P}$ (and hence any associated dense classes and G) can now be proper-class-sized rather than just set-sized, and so while $\mathbb{P} \subset \mathfrak{M}$, it is not in general true that $\mathbb{P} \in \mathfrak{M}$, and G is added as a *class* (from which it may be possible to construct new sets). The construction goes through largely the same as set forcing, with a few additional intricacies and features.[10]

Class forcing has some interesting properties not enjoyed by standard set forcing. A good example is that if we allow class forcing over L, then there are reals that we can add using class forcing that cannot be added by set forcings.[11] Further, using class forcing we can produce models that violate **ZFC**. For example, the partially ordered class $Col(\omega, On)$ (i.e. functions p from finite subsets of ω into On ordered by reverse inclusion) is (without further constraints) perfectly legitimate. But if we force with this poset we obtain a model $\mathfrak{M}[G]$ that satisfies **ZFC** as long as G is *not* allowed as a class predicate, as the first-order domains of $\mathfrak{M}$ and $\mathfrak{M}[G]$ are identical.[12] However, if we admit G as a predicate into the language, Replacement fails. This is because G codes a cofinal sequence from ω to $On^{\mathfrak{M}[G]}$ and there is no set in $\mathfrak{M}[G]$ corresponding to $On^{\mathfrak{M}[G]}$.[13] This is unlike the case of set forcing where $\mathfrak{M}[G]$ is guaranteed to satisfy **ZFC** if $\mathfrak{M}$ does. If **ZFC** preservation is desired[14] some care[15] is required in defining the relevant $\mathbb{P} \subseteq \mathfrak{M}$ to be used in forcing.

Class forcing thus introduces two additional challenges not posed by set forcing. Whilst we must provide an interpretation of any sets that get added, we must also explain what happens with a forcing that violates **ZFC**, and also how we should understand the addition of a class without the addition of any sets that could underpin the 'change' in classes.[16]

[10] One option here is to force over models of the form $L(A) = \bigcup\{L(A \cap V_\alpha) | \alpha \in On\}$. Any model (M, A) of **ZF** (where we include Replacement for formulas mentioning A) can be changed to a model of this form by expanding it to a model (M, A^*) where $A^* = \{\langle 0, x\rangle | x \in A\} \cup \{\langle 1, V_\alpha^M\rangle | \alpha \in On^M\}$. Details of this presentation are available in [Friedman, 2000], Chapter 2. A second (more recent) option is to proceed directly in a second-order set theory. See [Antos, 2018] for explanations of approaches of this method.

[11] This is a deep result of Jensen, see [Friedman, 2010], p. 559 for details.

[12] To see this, note that for any $\mathbb{P}$-name σ for this poset and for each condition p in the intersection of the transitive closure of σ with $\mathbb{P}$, $ran(p) \subseteq rank(\sigma)$. We then define the dense set $D = \{p \in \mathbb{P} | rank(\sigma) \in ran(p)\}$. D is then both dense and definable over $\mathfrak{M}$. Letting $\sigma^p = \{\tau^p | \exists q \in \mathbb{P}[\tau, q \in \sigma \wedge p \leqslant_{\mathbb{P}} q]\}$. We then have $\sigma^p = \sigma^G \in M$ whenever G is $\mathbb{P}$-generic over $\mathfrak{M}$ and $p \in D \cap G$, because p either extends or is incompatible with any condition in the transitive closure of σ. Hence, whenever G is $\mathbb{P}$-generic over $\mathfrak{M}$, they contain exactly the same first-order objects.

[13] For details, see [Holy et al., 2016].

[14] We shall argue later that for the purposes of talking about forcings over V, there is no a priori reason why **ZFC** preservation is especially desirable.

[15] The relevant conditions are *pretameness* and *tameness* of the partial order, corresponding respectively to preservation of Replacement and Power Set. See [Friedman, 2000] for details.

[16] A delicate issue here is exactly what theory our starting universe satisfies when performing a class forcing. [Friedman, 2000] is a thorough presentation, and studies class forcing over structures of the form (M, A), where A is first-order definable with parameters. However, we might also try to force over models of full Morse-Kelley class theory, see [Antos, 2018] for discussion. We will suppress these details for philosophical ease.

2 Looking at V through the Multiversist lens

We thus have two kinds of construction that seem to add external entities to models, and the datum that often set theorists use the terms 'V' and '$V[G]$' when performing a forcing construction. However, more must be done philosophically to create a genuine problem here. For example, why (given the Universist's 'predicament) can't she just insist that the use of the terms 'V' and '$V[G]$' should not be understood literally, or that forcing is simply not a legitimate construction? In this section we'll provide some support to the idea that there is a genuine problem to be dealt with here, and one way of motivating this problem is via constraints on the naturality of interpretation.

We will simply take it as read that the Universist should not outright reject forcing as a legitimate mathematical technique. Forcing is an essential part of the working set theorist's toolkit, and clearly represents a coherent body of study. We are also moved by the more open-ended desire to provide the Universist's opponent with the tools she requires to construct her objection, rather than just rejecting the grounds for the question outright.

Therefore throughout the paper we will analyse ways in which the Universist can interpret forcing talk. However just to begin setting things up, we should note that the Universist *can* provide interpretations of forcing that are *non-vacuous*. We discuss these approaches in detail §3, but two standard approaches are either to regard forcing statements as a façon de parler for speaking about Boolean-valued models or as concerned with countable transitive models of (a fragment of) **ZFC**.[17] Each method of interpretation, however, yields a model that does not commit the Universist to sets outside V.[18] This is pertinent, since one important use of forcing constructions has been to show *independence* results. Specifically, given a statement ϕ (such as the continuum hypothesis) we use forcing to construct a model of (a fragment of) $\mathbf{ZFC}+\neg\phi$, thereby showing that ϕ is not a consequence of (a fragment of) **ZFC** (or some extension thereof). If all we wish to do is show essentially number-theoretic facts about relative provability, then there is little pressure to accept that forcing requires anything more than a non-vacuous interpretation; all we need to do is show that there is *some model or other* that can serve as the interpretation of the forcing language and as a counterexample to the claim that $\mathbf{ZFC} \vdash \phi$, and the Universist can already do this.

In order for the existence of forcing to provide evidence against the Universist then, more needs to be said. In the rest of this section we will develop this problem. We will argue for the following claims:

(i) Forcing is more than just a tool for proving number-theoretic consistency facts,

[17] Indeed, this is the strategy of many set theory textbooks. For example the Boolean-valued model strategy is put forward by [Jech, 2002]:

> "As the properties of the generic extension can be described entirely withing the ground model, statements about $V[G]$ can be understood as statements in the ground model using the language of forcing." ([Jech, 2002], p. 201)

Whereas the countable transitive model strategy is the main strategy pursued in [Kunen, 2013] (see Ch. 4 §5).

[18] Koellner, for example, is sensitive to this point:

> "The advocate of the universe view is unmoved by these considerations [arising from forcing] since the models produced are not candidates for the universe of sets, the first because it is an object within the universe of sets, the second because it is a description of a class-size structure which is not even of the relevant type. The mere existence of the model-theory of forcing (something that is uncontroversially accepted by both parties) is not sufficient to secure the multiverse conception. Something more needs to be said." ([Koellner, 2013], pp. 18–19)

it is also useful for formulating axioms and proving theorems about large infinitary objects in V.

(ii) Given an acceptance of the claim that providing a 'natural' or 'transparent' interpretation of mathematical discourse is desirable, the Universist has additional constraints on her interpretation of forcing.

2.1 Proving theorems and formulating axioms

How is forcing useful for more than proving number-theoretic facts? As we'll see, forcing over V can be used in analysing the structure of large infinitary objects outside the domain of number theory. This is both with respect to proving theorems from *within* **ZFC**, but also for formulating axioms that go *beyond* **ZFC** (and proving theorems on their basis). As we'll discuss below, this puts some pressure on the claim that we might want *more* than just any old interpretation of the mathematics—since viewing V as though it were part of a multiversist framework facilitates insight concerning V, we might desire an interpretation of this talk that respects as much intuitive content of the mathematics as possible. Before we provide this philosophical analysis though, we provide a brief description of the mathematical landscape to show the importance of this way of thinking.

Proving theorems

Within **ZFC**, there are a wide number of questions concerning V that can be settled on the basis of considering forcing extensions. The rough strategy of such theorems is to show that if V has a forcing extension such that ϕ (for some particular ϕ) then some other sentence ψ holds of V (say by using absoluteness facts). For example:[19]

Theorem 1. [Baumgartner and Hajnal, 1973] $\omega_1 \longrightarrow (\alpha)^2_n$ for all finite n and countable α (i.e. For all finite n and countable α, every partition of the two-element subsets of ω_1 into a finite number of pieces has a homogeneous[20] set of order-type α).

The proof proceeds by finding a homogeneous set in a forcing extension where MA holds. This establishes that a certain tree from the ground model is non-well-founded in the extension. We then know, by the absoluteness of well-foundedness, that the tree is also non-well-founded in the ground model, establishing the theorem.

The theorem is broadly illustrative of how one can use the perspective of extensions to prove facts about the ground model. One moves to an extension where one has ensured the existence of objects of a certain desirable kind, used these to reason about objects in the ground model, and then inferred via absoluteness facts (e.g. Lévy-Shoenfield absoluteness) that the ground model must in fact have the relevant properties. Importantly here, theorems like the above are not straightforwardly about number theory and the possibility of proof, but rather concern genuine infinitary objects in the ground model.

A compendium of similar theorems is available in [Todorčević and Farah, 1995]. One further clear example bears mentioning here though. Since the formulation of CH (and realisation that it is independent from **ZFC**) set theorists have studied so-called *cardinal characteristics*; the study of uncountable infinite sets that may or may not be smaller than the continuum. The relationships of these infinite cardinals

[19] I am grateful to Andrés Caicedo for pointing out this example.

[20] Here, a homogeneous set is a subset X of ω_1 such that every 2-element subset of X is in the same member of the partition.

are well studied[21], but recently [Malliaris and Shelah, 2016] showed that two such characteristics (namely $\mathfrak{p}$ and $\mathfrak{t}$) that were previously thought to be likely separable in certain models were in fact equal. The proof is rather involved[22], but depends crucially on supposing for contradiction that $\mathfrak{p} < \mathfrak{t}$ in V, and then tracing out some consequences of this assumption (and finding a contradiction) in a forcing extension $V[G]$. Thus, situating V within a Forcing Multiversist framework (i.e. someone who thinks that no universe is maximal and we can *always* move to a forcing extension) allows us to solve important[23] and difficult set-theoretic questions on the basis of how V behaves with respect to its forcing extensions.

Thus, forcing extensions can function with respect to V in contemporary set theory somewhat like the historical situation with complex and real numbers. For several years, before the advent of interpretation of the complex numbers in the Euclidean plane, mathematicians were still using the algebraic properties of complex numbers to prove results about the reals with great success.[24] Similarly, extensions of V can be used to determine properties of V. It is the external *perspective* provided by the forcing extension that can facilitate a greater understanding of the ground model. The Universist then, might feel some pressure to interpret forcing over V—by looking at V from $V[G]$, she is able to see more.

Formulating axioms

The same is true also for theories *extending* **ZFC**. Especially interesting here is that forcing becomes a useful tool not just for *proving theorems*, but also *formulating axioms*. As we explain below, the flexibility afforded by extensions often provides us with additional resources for expressing axioms with interesting properties.

One way of asserting the existence of large cardinals (often quite strong), is through the use of elementary embeddings. The cardinals *measurable*, *strong*, *supercompact* (among others) are all naturally defined by positing the existence of elementary embeddings from V into transitive inner models. These represent *strong* axioms, pushing us (unlike smaller large cardinals) well beyond $V = L$. When defining a large cardinal through an embedding $j : \mathfrak{N} \longrightarrow \mathfrak{M}$, the strength of the embedding depends mainly on two parameters:

(i) The size of $\mathfrak{N}$ and $\mathfrak{M}$.

(ii) Where j sends the ordinals.

For instance, the minimal case for an embedding between proper class models; namely the existence of a non-trivial $j : L \longrightarrow L$ suffices to define the principle that "$0^\sharp$ exists". If we assume that $dom(j) = V$, we strengthen to the level of a measurable cardinal. We know that the existence of a non-trivial $j : V \longrightarrow V$ is

[21] Chichoń's diagram, for example, is a manifestation of the detailed knowledge we now have concerning how some of these characteristics behave in models of **ZFC**.

[22] I thank Jonathan Schilhan for patient explanation here.

[23] For instance, the result (unlike much of set theory) attracted some interest from the wider mathematical community, being the subject of a piece in *Quanta* magazine and receiving some attention from Tim Gowers.

[24] See, for example the remarks in [Painlevé, 1900] concerning real and complex analysis:

> "The natural development of this work soon led the geometers in their studies to embrace imaginary as well as real values of the variable. The theory of Taylor series, that of elliptic functions, the vast field of Cauchy analysis, caused a burst of productivity derived from this generalization. It came to appear that, between two truths of the real domain, the easiest and shortest path quite often passes through the complex domain."

inconsistent (*modulo* **ZF**) with AC.[25] Despite this we can study intermediate cardinals by modifying the properties of j and $\mathfrak{M}$. For example, we can use the following pair of definitions to strengthen the notion of measurable along the dimensions of (i) and (ii):

Definition 2. A cardinal κ is λ-*supercompact* iff it is the critical point of a non-trivial elementary embeddings $j : V \longrightarrow \mathfrak{M}$, such that $j(\kappa) > \lambda$ and ${}^{\lambda}\mathfrak{M} \subseteq \mathfrak{M}$ (i.e. $\mathfrak{M}$ is closed under λ-sequences).

Definition 3. A cardinal κ is *supercompact* iff it is λ-supercompact for all $\lambda \in On$.

The definition of supercompact uses the dimensions of (i) and (ii) to increase the strength of the embedding. We postulate a higher degree of similarity between V and $\mathfrak{M}$ (in terms of closure under λ-sequences for the relevant λ), and stipulate that j sends κ above λ.

Standard discussions of middling large cardinals proceed from this template. However, we can generalise the construction to *generic* embeddings. Given a forcing construction adding a generic G over a model $\mathfrak{N}$, a generic embedding is of the form $j : \mathfrak{N} \longrightarrow \mathfrak{M} \subseteq \mathfrak{N}[G]$. In other words, we begin to study embeddings from structures to inner models *of their forcing extensions*.

Recently, there has been an increased focus on such embeddings. Indeed, the study of generic embeddings has become widespread, as Foreman (in a *Handbook of Set Theory* article on generic embeddings) illustrates:[26]

> "The main aim of the chapter is to illustrate that there is a coherent theory here, that there are unifying fundamental ideas that occur frequently in many different contexts. These include master condition ideals, natural and induced ideals, disjointing, self-genericity, the role of diagonal unions for representing Boolean sums, good elementary substructures—the list is long." ([Foreman, 2010], p890)

Generic embeddings are thus useful for studying certain natural mathematical properties. Furthermore, the involvement of extensions in the consideration of embeddings provides an additional dimension in which we may vary the nature of the construction. Not only does the embedding depend upon the size of the domain and range of the embedding and where the ordinals are sent, but also on a third parameter:

(iii) The nature of the forcing required to define j.[27]

The fact that we have an extra dimension in which we can vary the structure of these kinds of embeddings makes them an intriguing subject matter. However, even more interesting is that the critical points of these axioms can be rather small. For example, we have:

Theorem 4. If I is an ω_2-saturated ideal on ω_1 and U is generic for the poset of I-positive sets, then in $V[U]$ the ultrapower $Ult(V, U)$ is well-founded and we get a map $j : V \longrightarrow \mathfrak{M} \subseteq V[U]$ with $crit(j) = \omega_1$ and $j(\omega_1) = \omega_2$.

[25] See [Kunen, 1971] for the result, and [Schindler, 2014] for a recent presentation.

[26] See also, [Foreman, 1986] for several key results, and [Foreman, 1998] for a more informal overview.

[27] For further exposition of this line of thinking, see [Foreman, 1998] and [Foreman, 2010].

Such an embedding from V to another model $\mathfrak{M}$ has ω_1 as its critical point, far below the size of a measurable cardinal.[28] Despite the smallness of the critical points, however, these embeddings have a significant amount of large cardinal strength.[29] Thus, these embeddings provide significant combinatorial power whilst facilitating proof concerning *small* uncountable sets.[30]

Moreover, it is not just through postulating the existence of embeddings *directly* between V and models within forcing extensions that allow us to formulate new axioms with significant large cardinal strength. Through analysing properties of ordinals in extensions, we can come to characterisations of new varieties of cardinal. A relatively recent development is the emergence of *virtual* large cardinals. Here, we postulate that a particular ordinal or initial segment of V has a certain large cardinal property *in a forcing extension* (as opposed to *internal* to V). Various generalisations of the standard large cardinal notions have been studied in this framework, including the notions of *virtually supercompact*, *virtually strongly compact*, *virtually strong*, *virtually Woodin*, and *virtually extendible*. As it turns out, the Kunen inconsistency does not hold for virtual embeddings (in the sense that one can have elementary embeddings $j : V_\lambda \to V_\lambda$ with λ far greater than the supremum of the criticial sequence), and so one even has a notion of *virtually rank-into-rank* without the usual restrictions on the size of the rank. Especially interesting is that these cardinals are consistent with $V = L$, in stark contrast to their 'non-virtual' counterparts. They thus provide an elegant hierarchy that helps to elucidate the large cardinal landscape between ineffable cardinals and $0^\sharp$.[31]

We will consider just one example to help see how these cardinals are defined:

Definition 5. [Schindler, 2000] A cardinal κ is *remarkable* iff in the $Col(\omega, <\kappa)$ forcing extension $V[G]$, for every regular $\lambda > \kappa$ there is a cardinal $\lambda_0 < \kappa$, λ_0 regular in V, and $j : H^V_{\lambda_0} \longrightarrow H^V_\lambda$ such that $crit(j) = \gamma$ and $j(\gamma) = \kappa$.

We are able to characterise the notion of a cardinal being remarkable if, when we collapse all cardinals less than κ to ω through forcing, in this $Col(\omega, <\kappa)$ extension $V[G]$, for every regular $\lambda > \kappa$ in $V[G]$ there is a V-regular cardinal $\lambda_0 < \kappa$ such that the heriditarily λ_0-sized sets elementarily embed into the hereditarily λ-sized sets. The definition turns out to be a characterisation of the notion of *virtual supercompactness*[32], and lies in consistency strength between a 1-iterable and 2-iterable cardinal. While not strong enough to push us outside $V = L$ (no known virtual cardinals can),

[28] We know, for instance, that $|\omega_1| \leqslant |\mathcal{P}(\omega)|$, making it accessible. To get an idea of the scale of the difference, if κ is measurable then it has to be an inaccessible limit of inaccessible cardinals.

[29] For example, the existence of both a saturated ideal on ω_1 (and associated generic embedding) and a measurable cardinal implies the existence of an inner model with a Woodin cardinal, whereas the consistency strength of a measurable cardinal is far below that of a single Woodin. See [Steel, 1996] for details.

[30] Foreman, for example, writes concerning generic embeddings:

> "The advantage of allowing the embeddings to be generic is that the critical points of the embeddings can be quite small, even as small as ω_1. For this reason they have many consequences for accessible cardinals, settling many classical questions of set theory." ([Foreman, 2010], p887)

.

[31] An excellent survey of the recent developments in virtual large cardinals mentioned here is available in [Gitman and Schindler, 2018]. Two particular uses of these cardinals are to study *Silver indiscernibles* in L and index the consistency strength of other kinds of virtual axioms. In fact, we could have spoken for longer about different kinds of virtual principle, such as *virtual forcing axioms*, but considerations of space prevent a presentation of the full picture. For an overview, references, and further reading the reader is once again directed to [Gitman and Schindler, 2018].

[32] See [Gitman and Schindler, 2018], p. 2.

it is substantially stronger than a weakly compact cardinal[33]. Thus, by studying how sets are embeddable *in the extension*, we are able to ascribe large cardinal properties to ordinals *in* V. The problem then for the Universist is that we are predicating a large cardinal property of κ, an object *in* V, but using resources from extensions to define what it is to have said large cardinal property.

The case of class forcing is somewhat trickier, since it has received less attention than set forcing. However, some uses have been found in the literature; for example certain generic embeddings can be defined using class-sized stationary tower forcing (on the assumption that V satisfies large cardinal properties).[34]

Moreover, one can show that class forcing facilitates the formalisation of different axioms that go beyond what can be captured by set forcing.

A good example here is:

Axiom 6. *The Inner Model Hypothesis.* Let ϕ be a parameter-free first-order sentence. By an *outer model* of a model $\mathfrak{M}$, we mean a model $\mathfrak{O}$ satisfying **ZFC** with the same ordinals as $\mathfrak{M}$, and such that $\mathfrak{M} \subseteq \mathfrak{O}$. Then the *Inner Model Hypothesis* for $\mathfrak{M}$ states that if ϕ is true in an inner model of an outer model of $\mathfrak{M}$, then ϕ is already true in an inner model of $\mathfrak{M}$.

The Inner Model Hypothesis, as proposed by [Friedman, 2006], is meant to apply to *arbitrary* width extensions of a model, as well as *arbitrary* inner models. It is therefore often stated as second-order assertions concerned with a countable transitive models $\mathfrak{M} = (M, \in, \mathcal{C}^{\mathfrak{M}})$ in some ambient universe (possibly V), where quantification over $\mathfrak{M}$ and its outer models is uncontroversial. Recently, however, Antos, Barton, and Friedman showed that by using infinitary logics to code satisfaction in outer models, and coding the infinitary logic using using proper-class-sized trees, one can formulate versions of the full IMH in a variant of **MK**.[35] Since we are interested in *forcing* extensions, we will restrict ourselves to considering the following two below restricted forms. Since we are just setting up the axioms at this stage, one can regard them for the moment as concerned with a countable transitive model $\mathfrak{M} = (M, \in, \mathcal{C}^{\mathfrak{M}})$ of some second-order set theory, in a wider second-order multiverse (possibly given by some ambient universe). Again, letting ϕ be parameter free we have:

Axiom 7. $\mathfrak{M}$ satisfies the *Set-Generic Inner Model Hypothesis* iff whenever a (first-order, parameter free) sentence ϕ holds in an inner model of a set forcing extension $\mathfrak{M}[G] = (M[G], \in, \mathcal{C}^{\mathfrak{M}[G]})$ of $\mathfrak{M} = (M, \in, \mathcal{C}^{\mathfrak{M}})$ (where $M[G]$ consists of the interpretations of set-names in V using G, and $\mathcal{C}^{\mathfrak{M}[G]}$ consists of the interpretations of class-names in $\mathcal{C}^{\mathfrak{M}}$ using G), then ϕ holds in an inner model of $\mathfrak{M}$.

[33] Weakly compact cardinals are so named in virtue of being characterisable through compactness properties on infinitary languages. They admit of a diverse number of equivalent characterisations. For details, see [Kanamori, 2009].

[34] The following example is taken from [Larson, 2004], §2.3, p. 59. Suppose that V contains a proper class of completely Jónsson cardinals. Letting $\mathbb{P}_\infty$ be the class tower forcing, and $G \subset \mathbb{P}_\infty$ be V-generic, and $V[G]$ be:

$$V[G] = \bigcup_{\alpha \in Ord} L(V_\alpha, G \cap V_\alpha)$$

there exists a generic embedding $j : V \longrightarrow V[G]$ such that for every $a \in \mathbb{P}_\infty$, $a \in G$ iff $j[\cup a] \in j(a)$. I thank Monroe Eskew for bringing this example to my attention.

[35] See [Antos et al., S] for discussion. One requires an additional assumption on the existence of isomorphisms, which Antos, Barton, and Friedman add to the second-order set theory. The proof does not use the full strength of class comprehension, but rather can be accomplished with Σ^1_1-Comprehension. However, Σ^1_1-Comprehension contradicts the usual IMH, and so in this framework the IMH is formalisable, but false. Antos, Barton, and Friedman remedy this by restricting to outer models which *also* satisfy Σ^1_1-Comprehension.

This admits of a natural generalisation:

Axiom 8. Again, let $\mathfrak{M} = (M, \in, \mathcal{C}^{\mathfrak{M}})$ be a **NBG** structure. The *Class-Generic Inner Model Hypothesis* is the claim that if a (first-order, parameter free) sentence ϕ holds in an inner model of a tame class forcing extension $\mathfrak{M}[G] = (M[G], \in, \mathcal{C}^{\mathfrak{M}[G]})$ of $\mathfrak{M} = (M, \in, \mathcal{C}^{\mathfrak{M}})$ (where $\mathfrak{M}[G]$ and $\mathcal{C}^{\mathfrak{M}[G]}$ are defined as above), then ϕ holds in an inner model of $\mathfrak{M}$.

If $\mathfrak{M}$ satisfies the Set-Generic Inner Model Hypothesis we get some consequences:

Fact 9. If $\mathfrak{M}$ satisfies Set-Generic Inner Model Hypothesis, then $\mathfrak{M} \models V \neq L$.[36]

However, though the Set-Generic Inner Model Hypothesis is sufficient to get us a certain richness of inner models (enough to break $V = L$) we get more if we allow class forcings. This is brought out in the following:

Fact 10. Assuming the consistency of an model $\mathfrak{M}$ with a $(V_\kappa)^{\mathfrak{M}} \prec \mathfrak{M}$, there is a model satisfying the Set-Generic Inner Model Hypothesis that does not satisfy the Class-Generic Inner Model Hypothesis.[37]

It thus seems that class forcing is able to aid in the facilitation of formulating axioms that substantially go beyond those that are formalisable using set forcings.[38] Of course, we have thus far formulated these axioms as about countable transitive models, which are not the 'absolute' universe. We will discuss the merits of different formulations of extension-talk involving V later, but for now we note that these versions *can* be formulated by the Universist, assuming that she can use **NBG** over V. In each case, we can use the class theory to quantify over inner models (that may not be definable) whilst interpreting truth in an inner model of a forcing extensions by what sentences (in the second order set theory) are forced by particular conditions. We *can* (given some class theory) thus formulate these axioms as about V. Viewing V as situated in a multiverse framework (suitably coded), may then yield insight.

[36] We can see this by the following:

Proof. Assume $\mathfrak{M} \models V = L$ and that $\mathfrak{M}$ satisfies the Set-Generic Inner Model Hypothesis. Then there is an inner model of an outer model in which $V = L$ is false (the addition of a single Cohen real x over L to $L[x]$ will suffice, with the relevant inner model simply being the forcing extension $L[x]$). By the Set-Generic Inner Model Hypothesis there is an inner model of L in which $V \neq L$. But L is the smallest inner model, and so $\mathfrak{M} \models V = L$ and $\mathfrak{M} \models V \neq L, \bot$. □

[37] I am grateful to Sy Friedman and Monroe Eskew for discussion of the following:

Proof. Take a model $\mathfrak{M}$ of $V = L$ containing a $V_\kappa^{\mathfrak{M}} = L_\kappa^{\mathfrak{M}} \prec V^{\mathfrak{M}} = L^{\mathfrak{M}}$. We work from the perspective of $\mathfrak{M}$. For any particular β, let $Col(\omega, \beta)$ be the Lévy collapse of β to ω, and let G be generic for $Col(\omega, \kappa)$. We claim that $(L[G], Def^L[G])$ satisfies the Set-Generic Inner Model Hypothesis. It suffices to show that if a sentence holds after forcing with $Col(\omega, \lambda)$ for some λ, then this λ can be chosen to be less than κ. This is because any set-forcing can be absorbed into $Col(\omega, \lambda)$ for some λ (see here [Cummings, 2010] (§14) for technical explanation) and any two $Col(\omega, \lambda)$-generic extensions satisfy the same sentences.

Now if ϕ holds after forcing with $Col(\omega, \lambda)$ for some λ, then as L_κ is elementary in L, ϕ also holds after forcing over L_κ with $Col(\omega, \lambda')$ for some $\lambda' < \kappa$. But if $G(\lambda')$ is generic over L_κ for $Col(\omega, \lambda')$, then $L_\kappa[G(\lambda')]$ is elementary in $L[G(\lambda')]$ as $Col(\omega, \lambda')$ is a forcing of size less than κ, using the fact that κ is a cardinal of L. So ϕ holds after forcing with $Col(\omega, \lambda')$ for some $\lambda' < \kappa$, as desired.

However, $(L[G], Def^L[G])$ does not satisfy the Class-Generic Inner Model Hypothesis because every real in $L[G]$ is set-generic over L, and by Jensen coding any model of the Class-Generic Inner Model Hypothesis must have reals which are not set-generic over L. □

[38] Since we are interested in what is *possible* for a Universist, we will say little about motivation for these axioms here. However, one way of motivating these sorts of axioms is as *absoluteness principles*. Effectively, they say that V already has witnesses for ways things 'could' be (in the sense of possibility in coded extensions).

2.2 The Naturalness Constraints

Above, we saw how situating V in a multiverse of universes obtained by forcing facilitates the proving of theorems and formulation of axioms beyond the merely number-theoretic. However, we also noted that the Universist could at least interpret this talk *non-vacuously*. How might we then combine the observation that forcing is more than a tool for proving independence into a genuine problem for the Universist?

The key notion will be the *naturalness* of interpretation provided. As we have just argued, the proving of theorems and formulation of axioms in the present context relates to more than the mere existence of models to required to prove relative consistency proofs. Forcing allows us to prove that certain uncountable cardinals are in fact equal (e.g. $\mathfrak{p} = \mathfrak{t}$), posit the existence of and study large cardinals (e.g. remarkable cardinals), and formulate axioms with consequences for the existence of real numbers and accessible cardinals (e.g. generic embeddings and variants of the inner model hypothesis).

We might then press the following idea: It is not enough to merely give *some* model or other in providing an interpretation of forcing. Since it seems that set-theorists can use extensions of V to reason about V, and that when performing this reasoning the relevant extensions *appear* to be bona fide universes of set, we should want whatever interpretation we prescribe to be as faithful as possible to the set theorist's intuitive thinking. Perhaps under such constraints the Universist's interpretations *will* come out as deficient?

There are two immediate challenges to this line of attack:

(1.) The notion of what counts as a 'natural' or 'faithful' interpretation is imprecise.

(2.) It is unclear why the Universist cannot just reject this requirement on naturalness outright.

Both challenges can be answered, at least partially. Regarding (1.): While it is true to say that what the notion of being a 'natural', 'faithful', or 'semantically transparent' interpretation comes down to is going to be a difficult question to answer (in the sense of providing necessary and sufficient conditions) this does not mean that we have *zero* grip on the notion. In fact, the following all seem to be legitimate and precise constraints on ways in which an interpretation can be more or less natural:

(1.) (The Facetious Constraint.) The interpretation of 'V' could refer to V, G to an actual generic outside V, and $V[G]$ to a *literal* extension of V.

(2.) The interpretation of 'V' in the construction could be V itself.

(3.) More minimally, the interpretation of 'V' in the construction could satisfy the same first-order sentences as V.

(4.) One or both of the interpretations of 'V' and '$V[G]$' could be well-founded, and hence admit of an absolute notion of being formed through transfinite iteration of a powerset-like operation.

(5.) The interpretations of 'V' and '$V[G]$' could contain uncountable sets.

(6.) Each of 'V' and '$V[G]$' could contain all the ordinals.

(7.) The structures denoted by each of 'V' and '$V[G]$' could be two-valued.

(8.) The movement between the interpretation of 'V' and '$V[G]$' could be 'transparent', in the sense that whatever is denoted by '$V[G]$' really is obtainable by the usual forcing idea of the addition of a generic to whatever is denoted by 'V'.

(9.) Steps in proofs that use forcing constructions could be interpreted with the minimal amount of change, so additional or different steps do not need to be made to keep the proof in line with the interpretation.

We will refer to the project of trying to satisfy as many of the above desiderata as possible as satisfying the *Naturalness Constraints*. While we certainly do not contend that the list is complete, we think that the ease of providing such a preliminary list shows that we have some grip on the notion of what counts as more or less natural as an interpretation. The Facetious Constraint is so-named because it obviously cannot be satisfied by the Universist, but as we shall see shortly is helpful in clarifying the dialectic.

Regarding (2.): The question of whether or not the Universist should accept the Naturalness Constraints to *any* degree is going to be a difficult one. However, some points can be made in the favour of acceptance.

First, we might accept the Naturalness Constraints just for the sake of examining the opponents challenge on the strongest grounds possible. If the Universist can provide an interpretation that respects as many of the Naturalness Constraints as possible, this represents a more robust conclusion and so is philosophically desirable anyway.

Second, responding to the Naturalness Constraints is particularly *dialectically* effective given that some opponents to Universism seem to subscribe to similar ideas. Consider the following passages from Hamkins (an ardent anti-Universist):

> "This abundance of set-theoretic possibilities poses a serious difficulty for the universe view, for if one holds that there is a single absolute background concept of set, then one must explain or explain away as imaginary all of the alternative universes that set theorists seem to have constructed. This seems a difficult task, for we have a robust experience in those worlds, and they appear fully set theoretic to us. The multiverse view, in contrast, explains this experience by embracing them as real, filling out the vision hinted at in our mathematical experience, that there is an abundance of set-theoretic worlds into which our mathematical tools have allowed us to glimpse." ([Hamkins, 2012], p. 418)

and

> "...a set theorist with the universe view can insist on an absolute background universe V, regarding all forcing extensions and other models as curious complex simulations within it. (I have personally witnessed the necessary contortions for class forcing.) Such a perspective may be entirely self-consistent, and I am not arguing that the universe view is incoherent, but rather, my point is that if one regards all outer models of the universe as merely simulated inside it via complex formalisms, one may miss out on insights that could arise from the simpler philosophical attitude taking them as fully real." ([Hamkins, 2012], p. 426)

Part of Hamkins' point, it seems, is that while he acknowledges that the Universist *can* interpret forcing constructions, the interpretation provided is oftentimes *unnatural* requiring "contortions" and not respecting the "mathematical experience".

Better instead (he argues) to take a forcing multiversist position on which extensions are uncontroversially and easily available (indeed, the forcing multiversist immediately satisfies the Facetious Constraint). If the Universist can provide a response that satisfies as many of the Naturalness Constraints as possible, she would thus provide the strongest possible response to criticisms in the style of Hamkins.

Thirdly, the Naturalness Constraints chime well with things the Universist might be tempted to say anyway. A standard (but by no means the only) response one often receives when asking why one might hold the Universist picture is that it seems, given the iterative conception of set, that one is defining a unique structure. One simply iterates the powerset operation through the ordinals, thereby specifying V. Now, while the significance or effectiveness of such an argument is highly debatable, it nonetheless bears mentioning that the line of thinking relies on something like the Naturalness Constraints: It certainly *seems* as though we talk about a unique structure when we begin to study set theory through the iterative conception, and so the most *natural* interpretation (the Universist might argue) is that we do so. (Note that this remark (and the current debate) would apply equally well to theorists who do not hold that the universe has a definite height, but are nonetheless tempted by the idea that the powerset operation is definite.) In short, the Universist believes that there is an absolute universe V in which the greater-than-first-order notions of well-foundedness, uncountability, containing all the ordinals, etc. can be computed. An interpretation in which we have more of these object-language notions absolute between the interpretation in question and V (understood as the context for the meta-theory) is thus more natural than one in which fewer notions are absolute.[39] She is thus involved in a delicate balancing act in which she tries to find an interpretation that is as 'V-like' as possible, whilst keeping forcing as about a genuine extension of the interpretation by generic filters.

We thus believe that the Universist has at least some reason to try and provide an interpretation of forcing that respects the Naturalness Constraints. From now on we shall simply take it as an assumption that she would like to satisfy as much of the Naturalness Constraints as possible, whilst acknowledging that this is an area that merits further analysis.

As we shall see, extant interpretations of forcing constructions reveal that the Universist can satisfy a good deal of the Naturalness Constraints, although the various interpretations behave differently. First, however, let's examine a quick example to see how the Naturalness Constraints might exclude certain interpretations as fully satisfactory. Both remarkable cardinals and generic embeddings in fact admit of formulations internal to V, without mentioning forcing at all. Remarkable cardinals, for example, can be formulated in terms of embeddings between countable models and certain sets of particular hereditary cardinalities.[40] Generic embeddings can likewise be formulated by an internal to V condition, usually in terms of the existence of particular ideals.[41] One might then just regard the formulation of remarkable car-

[39] This is not to say that keeping the distinction between object-language and meta-theory isn't important, and that significant mathematical insights can't be gained by working in a non-standard meta-theory; a good example here being [Hamkins and Yang, 2013]. The point is rather that the Universist *believes* in the existence of these absolute notions, and thereby believes some contexts are more 'V-like' than others.

[40] See here [Schindler, 2001].

[41] Some remarks of Foreman are pertinent here:

> "The language of ideals, together with the mechanics of forcing provide the same kind of vehicle for stating generalized large cardinal axioms in the language of set theory. Assuming the existence of a proper class of Woodin cardinals, Burke's Proposition...shows that every countably complete ideal is pre-precipitous. More directly: the existence of an elementary embedding $j : V \longrightarrow \mathfrak{M} \subseteq V[G]$ where $G \subseteq \mathbb{P}$ is generic and $j"\lambda \in \mathfrak{M}$ is easily

dinals and generic embeddings as a façon de parler for ascribing these internal-to-V properties to certain objects.

Given the Naturalness Constraints, this strategy seems to perform relatively poorly. In particular, whilst there is no denying the mathematical interest of these equivalent characterisations, as an interpretation of the relevant forcing discourse there is no actual interpretation given to the forcing *itself*, no model *similar* to V is extended, and there is no *actual forcing* occurring at all. In this way, the interpretation does not seem very natural and somewhat contorted; part of the insight gained from having these axioms is facilitating the study of objects of V 'as if' they were situated in a forcing multiversist framework.[42] Often the most natural context in which to define and study them (and the relevant associated objects internal to V) is the forcing perspective. An interpretation of forcing which can make use of this perspective thus gains more insight than one which cannot.

3 Available interpretations

Thus far we've argued that (1.) The Universist has reason to want to interpret forcing talk over V (in order to study the objects of V) and (2.) Taking the Universist's opponent's complaints seriously depends on holding that some degree of 'naturalness in interpretation' is desirable. We will now analyse ways that we might code talk concerning forcing extensions, and see how they perform with respect to the Naturalness Constraints. We'll argue that different interpretations sacrifice different aspects of the Constraints, but that some are more serious than others. In the end we'll see that the so-called *Boolean ultrapower map* is especially natural for a certain class of forcings, and the *countable transitive model strategy* provides a method that satisfies many of the Naturalness Constraints. We'll also see that certain natural forcing contexts can be obtained through the use of large cardinals.

3.1 The forcing relation

We begin with a discussion of how we might try to capture claims concerning forcing extensions *syntactically* by defining a relation that captures the consequences of extensions without actually committing to the existence of any models. For forcing, this can be done by defining a so-called *forcing relation*. Roughly put, letting

> seen to be equivalent to the existence of a $\mathbb{P}$-term for an ultrafilter $\dot{U} \subseteq \mathcal{P}(\mathcal{P}(\lambda))^V$ is normal for regressive functions in V and fine and is such that there is no descending ω-sequence of U-equivalence classes of functions from V. The idea of an *induced ideal* allows us to restate this combinatorially as a normal, fine, precipitous ideal I on $\mathcal{P}(\lambda)$ such that the quotient algebra $\mathcal{P}(\mathcal{P}(\lambda))/I$ inherits some of the properties of the original partial ordering $\mathbb{P}$. Finally, moving along the "F" axis [the nature of forcing required to define j] in the direction of greater strength, the saturation properties of ideals play exactly the same role for generalized large cardinals as ultrafilters do for conventional large cardinals." ([Foreman, 2010], p1128)

[42] An anonymous reviewer helpfully points out that this is somewhat similar to the case where we consider large cardinals of measurables and stronger as defined through elementary embeddings $j : V \to \mathfrak{M}$. For example, a cardinal κ is measurable iff it is the critical point of an elementary embedding $j : V \to \mathfrak{M}$, and a cardinal κ is strong iff for every θ there is a $j : V \to \mathfrak{M}$ with V_θ contained in $\mathfrak{M}$. Since these formulations quantify over j, V, and $\mathfrak{M}$ they look second-order, but they also have first-order characterisations in terms of measures and extenders, of a much more complicated nature. The embedding characterizations are robust and philosophically pleasing, whereas the first-order characterizations are tedious and less enlightening.

$\mathbb{P} = \langle P, \leqslant_{\mathbb{P}} \rangle$ be a forcing poset, $p \in P$, and ϕ be in the forcing language for $\mathbb{P}$,[43], we can define a relation $p \Vdash^*_{\mathbb{P}}$ recursively.[44] One can then verify:

(1) If $\phi_1, ..., \phi_n \vdash \psi$ and $p \Vdash^*_{\mathbb{P}} \phi_i$ for each i, then $p \Vdash^*_{\mathbb{P}} \psi$.

(2) $p \Vdash^*_{\mathbb{P}} \phi$ for every axiom of **ZFC**.

(3) If $\phi(x_1, ..., x_n)$ is a formula known to be absolute for transitive models, then for every p and all sets $a_1, ..a_n$; $p \Vdash^*_{\mathbb{P}} \phi(\check{a}_1, ..., \check{a}_n)$ iff $1_{\mathbb{P}} \Vdash^*_{\mathbb{P}} \phi(\check{a}_1, ..., \check{a}_n)$ iff $\phi(a_1, ..., a_n)$ is true in V.

Essentially, $\Vdash^*_{\mathbb{P}}$ lets us talk about what would be satisfied in the extension $V[G]$ by analysing what sentences conditions $p \in P$ force. In particular, if we can show that there is a $p \in P$ such that $p \Vdash^*_{\mathbb{P}} \phi$, we can behave and talk as if such a forcing extension exists. By (3), any theorem proved 'in $V[G]$' will be verified by the check names and hence by specific sets in V. Similarly, if we wish to formulate an axiom about V using a forcing extension, we can do so by finding a p that forces the required sentence about objects in the ideal extension.

The use of the forcing relation is absolutely fine for relative consistency proofs. We know that if we can find a $p \in \mathbb{P}$ such that $p \Vdash^*_{\mathbb{P}} \phi$, then we cannot prove $\neg\phi$ (assuming the consistency of **ZFC**). There are, however, several problems with this approach to interpreting forcing talk.

First, there is a problem of scope. The difficulty lies in the fact that the forcing relation need not be definable when the forcing poset in question is *proper-class-sized*. For example, consider the following forcing:

Definition 11. Let $\mathfrak{M}$ be a model for **ZFC**. Then the *Friedman* poset (denoted by '$\mathbb{F}^{\mathfrak{M}}$') is a partial order of conditions $p = \langle d_p, e_p, f_p \rangle$ such that:

(i) d_p is a finite subset of ω.

(ii) e_p is a binary acyclic relation on d_p.

(iii) f_p is an injective function with $dom(f_p) \in \{\varnothing, d_p\}$ and $ran(f_p) \subseteq \mathfrak{M}$.

(iv) If $dom(f_p) = d_p$ and $i, j \in d_p$, then ie_pj iff $f_p(i) \in f_p(j)$.

(v) The ordering on $\mathbb{F}^{\mathfrak{M}}$ is given by:

$$p \leqslant_{\mathbb{F}^{\mathfrak{M}}} q \leftrightarrow d_q \subseteq d_p \wedge e_p \cap (d_q \times d_q) = e_q \wedge f_q \subseteq f_p.$$

This defines a proper-class-sized partial order as the individual f_p include every function from some finite subset of ω to a (sub)set of $\mathfrak{M}$, and hence there are proper-class-many such ordered triples (relative to $\mathfrak{M}$). The partial order adds a bijection $F_{\mathbb{F}}$ between ω and M, and a relation $E_{\mathbb{F}} \in \mathfrak{M}[G]$ such that $\langle \omega, E_{\mathbb{F}} \rangle$ and $\langle M, \in \rangle$ are isomorphic. If the forcing relation for $\mathbb{F}$ were definable, $\mathfrak{M}$ would then have access to its own truth definition (contradicting Tarski's Theorem).[45]

Thus we have:

Theorem 12. [Holy et al., 2016] (attributed to Friedman) $\Vdash^*_{\mathbb{F}}$ is not uniformly definable for $\mathbb{F}$.

[43]The *forcing language of* $\mathbb{P}$ is the collection of all formulas that can be formed by the usual logical operators from the language $\mathscr{L}_{\in}$ combined with a constant symbol for every name in $V^{\mathbb{P}}$ (the $\mathbb{P}$-names).

[44]See [Kunen, 2013] for details of the forcing relation and verification of the relevant proofs.

[45]For the details of the proof, and further discussion of the Truth and Definability lemmas in context of class forcing, see [Holy et al., 2016].

Hence, there are forcings for which there is no definition of the forcing relation in the ground model. This is true, despite the fact that $\mathbb{F}^{\mathfrak{M}}$ itself is definable over $\mathfrak{M}$. If we wish to use $\mathbb{F}^V$ in proving facts about V then, we cannot do so through consideration of a forcing relation in V. Since well-behaved forcings (so called 'tame' and 'pretame' class forcings)[46] have definable forcing relations, we know that any extension using $\mathbb{F}$ must violate **ZFC**.[47] One may feel that this provides a response: we should not consider such forcing arguments as legitimate for proving facts about V because the resulting 'model' we are trying to talk about is pathological—it is non-**ZFC**-preserving.

We hold that the Universist should have little truck with such a claim. It is true that the resulting extensions are pathological in violating **ZFC**. However, we should note that (from the Universist's perspective) the whole enterprise with which we are engaged (namely considering extensions of V) is somewhat pathological: such sets cannot *literally* exist. We are rather trying to code in V the effects of viewing V as part of a forcing multiversist framework despite the fact that the relevant extensions do not exist (strictly speaking). Thus, there seems to be no objection to considering models where, say, there is a bijection between ω and V (as is the case when forcing using $\mathbb{F}^V$). If consideration of such 'pathologies' has interesting consequences (such as the study of how truth predicates might behave) within V then there seems little a priori reason to prohibit their examination.

One might simply *postulate* that a definition can be given for $\mathbb{F}^V$ and other class-sized partial orders with non-first-order definable forcing relations. Such definitions could not be first-order, but interestingly such a hypothesis fits naturally in the space of second-order set theories between **NBG** and **MK**. [Gitman et al., 2017] showed that the hypothesis that every class partial order has a forcing relation is equivalent to the principle that transfinite recursions of class relations for ordinal length are legitimate (so called 'ETR_{Ord}'). Whether this issue of scope can be assuaged thus depends precisely on one's attitude to second-order class theory. Presumably the Universist who wishes to try and interpret class forcing over V has some sympathy with the use of second-order resources (since the classes added are often non-first-order definable), and so we set the problem of scope aside.

Regardless, the use of $\Vdash^*_{\mathbb{P}}$ fares reasonably poorly with respect to the Naturalness Constraints. For many cases of forcing, reasoning syntactically about the forcing relation does not accord with the phenomenological character of the set theorist's reasoning. They wish to reason about sets which can be combinatorially manipulated, embeddings which move ordinals, and so forth. Thus, while 'V' is able to denote V in this interpretation of forcing, there are no models involved (two-valued or otherwise), there is no actual forcing occurring, and no model which is actually extended. This issue is brought out clearly when we consider generic embeddings. These provide us with quintessentially combinatorial kinds of reasoning; we want to see what ordinals are moved by j (and where) and what the structure of the remaining sets looks like given the existence of j. Here, however, the reasoning is fully syntactic; we analyse which formulas particular $p \in \mathbb{P}$ force, and so are explicitly *not* working with sets in the above manner. In this way, the use of the forcing relation, though both useful and of mathematical interest, leaves some philosophical issues untouched.

[46] Pretameness implies the preservation of Replacement in a class forcing, and tameness additionally requires that the forcing preserve the Power Set Axiom. For details, see [Friedman, 2000].

[47] See [Holy et al., 2016].

3.2 Boolean-valued models

Next, we consider the use of Boolean-valued models.[48] Starting with a forcing poset $\mathbb{P}$, we can find a separative[49] partial order $\mathbb{Q}$, equivalent to $\mathbb{P}$ for forcing, and a (unique up to isomorphism) *Boolean completion* of $\mathbb{Q}$ (denoted by '$\mathbb{B}(\mathbb{Q})$').[50,51] We then consider the class of $\mathbb{B}(\mathbb{P})$-names (denoted by '$V^{\mathbb{B}(\mathbb{P})}$'), assign values from $\mathbb{B}(\mathbb{P})$ to atomic relations between them, and provide an inductive definition for the quantifiers.[52]

It is then routine to show that $V^{\mathbb{B}(\mathbb{P})}$ is a Boolean-valued model of **ZFC**. In particular every axiom (and hence every theorem) of **ZFC** has Boolean-value $1_{\mathbb{B}(\mathbb{P})}$ in $V^{\mathbb{B}(\mathbb{P})}$. Moreover, for the purposes of consistency proofs, we know that if we can assign ϕ a Boolean-value greater than $0_{\mathbb{B}(\mathbb{P})}$, then $\neg\phi$ is not a consequence of **ZFC** (as if $\neg\phi$ is a consequence of **ZFC**, then ϕ receives Boolean value $0_{\mathbb{B}(\mathbb{P})}$). In fact, an assignment of a Boolean value greater than $0_{\mathbb{B}(\mathbb{P})}$ to ϕ exactly mimics the satisfaction of ϕ in some $V[G]$, for V-generic G.

Thus, by discussing the Boolean-valued model $V^{\mathbb{B}(\mathbb{P})}$, we are able to capture the intra-V content of talking about set forcing extensions of V. For example, suppose that we wish to show that the satisfaction of ϕ in a set forcing extension by $G \subseteq \mathbb{P} \in V$ has consequence ψ within V. We can then take the Boolean completion $\mathbb{B}(\mathbb{P})$, and show that we can assign ϕ Boolean-value greater than $0_{\mathbb{B}(\mathbb{P})}$ in $V^{\mathbb{B}(\mathbb{P})}$. By tracing the Boolean-values back to V, we then know that V satisfies ψ.

There are several problems with the use of Boolean-valued models, however. It has two particular limitations when it comes to forcing, despite its ability to capture satisfaction in forcing extensions.

The first is again a problem of scope. It is unclear how to interpret class forcing on the present approach. For, in class forcing, the relevant partial order $\mathbb{P}$ is proper-class-sized, and hence unbounded in the V_α. When defining the Boolean completion $\mathbb{B}(\mathbb{P})$ we then encounter a difficulty. The usual method for defining a Boolean completion is to find a separative partial order equivalent to $\mathbb{P}$ for forcing (known as the separative quotient), and embed it into a Boolean algebra[53]. Effectively, we add a bottom element and the required suprema to form $\mathbb{B}(\mathbb{P})$.[54]

In the present context, however, it is unclear that this can always be done. Since the partial order always goes to the 'top' of V, one can not always assume that there will be space to add a bottom element and suprema. This is not to say that it is *impossible* to provide a Boolean completion for class forcings; as it turns out a class partial order has a class Boolean-completion in a model of **MK** precisely when all antichains are at most set-sized (known as the Ord-chain condition).[55] This does show, however, that the kinds of class forcings one can interpret using this technique are rather restricted.

[48] The Boolean-valued approach was developed by Scott and Solovay, with additional contributions by Vopěnka (among others). See [Smullyan and Fitting, 1996], p. 273 for historical details and references.

[49] A partial order $\mathbb{P} = (P, <_{\mathbb{P}})$ is *separative* iff for all $p, q \in P$, if $p \not\leq_{\mathbb{P}} q$ then there exists an $r \leqslant_{\mathbb{P}} p$ that is incompatible with q.

[50] For details of Boolean algebras (from which our presentation is derived) see [Jech, 2002], Chapter 7. A discussion of Boolean completions is available in ibid. Chapter 14.

[51] We will (mildly) abuse notation and use $\mathbb{B}(\mathbb{P})$ to refer to the relevant Boolean completion even when $\mathbb{P}$ is not separative (i.e. the Boolean completion obtained from a separative partial order $\mathbb{Q}$, such that $\mathbb{Q}$ is equivalent to $\mathbb{P}$ for forcing).

[52] See here, [Jech, 2002], Ch. 14.

[53] More formally, for any set-sized partial order $\mathbb{P}$, there is a Boolean algebra $\mathbb{B}(\mathbb{P})$ and an embedding $e : \mathbb{P} \longrightarrow \mathbb{B}(\mathbb{P})^+$ (where $\mathbb{B}(\mathbb{P})^+$ is the set of non-zero elements of $\mathbb{B}(\mathbb{P})$) such that for $p, q \in \mathbb{P}$: (i) if $p \leqslant_{\mathbb{P}} q$, then $e(p) \leqslant_{\mathbb{B}(\mathbb{P})} e(q)$, (ii) p and q are compatible iff $e(p) \wedge e(q)$, and (iii) $\{e(p) | p \in \mathbb{P}\}$ is dense in $\mathbb{B}(\mathbb{P})$.

[54] For the full details, see [Jech, 2002], Chapter 14.

[55] See [Holy et al., 2016] and [Holy et al., 2018] for the result (attributed to Hamkins).

However, putting aside the questions of scope here, we note that even for the forcings that can be interpreted this way, the Naturalness Constraints are left somewhat neglected. This is for similar reasons to the forcing relation. Though the approach is now somewhat model-theoretic, the two-valued aspect of the set theorist's reasoning is missed. When we reason with $V[G]$ it 'looks-like' we reason with a standard set-theoretic model, and the Boolean valued model $V^{\mathbb{B}}$ does not have this flavour—rather it looks like we reason with Boolean-valued 'probabilistic' sets (since, one can informally think of a Boolean-valued model $V^{\mathbb{B}}$ as assigning 'probabilities' from $\mathbb{B}$ to membership and equality). Moreover, again no model is *actually* extended in considering these constructions. To accord more fully with the Naturalness Constraints, we need a two-valued model that looks a lot like V, and also gets extended in the relevant construction.

3.3 Boolean ultrapowers and quotient structures

There are, however, ways of modifying Boolean-valued models to proper-class-sized two-valued structures. We provide an informal description of the technique to highlight philosophical issues, but it is developed in detail in [Hamkins and Seabold, 2012] (and referred to in Hamkins' paper on the multiverse [Hamkins, 2012]). Importantly, the method provides a way of finding models internal to V that bear forcing relationships to one another. More formally, one can prove:

Theorem 13. [Hamkins and Seabold, 2012] *The Naturalist Account of Forcing*. If V is the universe of set theory and $\mathbb{B}$ is a notion of forcing, then there is in V a definable class model of the theory expressing what it means to be a forcing extension of V. Specifically, in the forcing language with $\in$, constant symbols $\check{x}$ for every $x \in V$, a predicate symbol $\check{V}$ to represent V as a ground model, and a constant symbol $\mathring{G}$, the theory asserts:

(1) The full elementary diagram of V, relativised to the predicate $\check{V}$, using the constant symbols for elements of V.

(2) The assertion that $\check{V}$ is a transitive proper class in the (new) universe.

(3) The assertion that $\mathring{G}$ is a $\check{V}$-generic ultrafilter on $\check{\mathbb{B}}$.

(4) The assertion that the new universe is $\check{V}[\mathring{G}]$, and **ZFC** holds there.

Hamkins and Seabold achieve this is by first taking an ultrafilter U on the relevant Boolean-algebra $\mathbb{B}$ (for convenience sake, we now drop the notation $\mathbb{B}(\mathbb{P})$) and constructing a particular ultrapower embedding j_U (the so-called Boolean ultrapower map) between V and an inner model $\check{V}_U$. When we then form the *quotient structure* $V^{\mathbb{B}}/U$ of $V^{\mathbb{B}}$ (formed by taking the standard quotient structure), we find an interesting relationship between $\check{V}_U$ and $V^{\mathbb{B}}/U$: $V^{\mathbb{B}}/U$ is precisely the forcing extension of $\check{V}_U$ by U. One can verify that $V^{\mathbb{B}}/U \models \mathbf{ZFC}$ and also that if ϕ has Boolean-value greater than $0_{\mathbb{B}}$ in $V^{\mathbb{B}}$, then $V^{\mathbb{B}}/U \models \phi$.

Importantly, there is no need for the ultrafilter U to be V-generic here, and hence U can perfectly well be in V. In fact, when one constructs the Boolean ultrapower over some model of set theory $\mathfrak{M} = (M, E)$, the claim that U is $\mathfrak{M}$-generic is equivalent to the Boolean ultrapower j_U being trivial (i.e. letting E_U be the 'membership' relation defined by the Boolean ultrapower, j_U is an isomorphism between $\mathfrak{M}$ and $(\check{M}_U, E_U)$).

In the case where j_U is non-trivial on V, we map V to a *subclass* of itself (much as we do with a measurable cardinal embedding). Since $\check{V}_U$ is not the whole of V when

U is in V (and hence not V-generic), it is possible for a set external to $\check{V}_U$ to be our generic for $\check{V}_U$. Our interpretation of forcing might then be as follows. We note that $\check{V}_U$, whilst not isomorphic to V, nonetheless looks a *lot* like V; it is a proper-class-sized elementary extension of V. Instead of using the Boolean-valued model $V^{\mathbb{B}}$ (with its attendant difficulties regarding classicality and combinatorial properties), we could interpret our use of the term 'V' and '$V[G]$' as $\check{V}_U$ and $V^{\mathbb{B}}/U$ respectively. We would then perform far better with respect to the earlier outline Naturalness Constraints, we have a two-valued structures, both proper-class sized, and there is actual model-theoretic extension of these structures occurring.

There are, however, several problems with this approach. We should first note that the use of Boolean ultrapowers and quotient structures suffers from the same problem of scope as the Boolean-valued model approach—since it depends on finding Boolean completions we will not be able to use the method for certain kinds of class forcing. We have, however, moved to an interpretation on which the reasoning is both two-valued and combinatorial, and on which a model very similar to V is extended.

However, this construction can often come at the price of well-foundedness. Observe that if the Boolean ultrapower map is to be well-founded, it must elementarily embed V into an inner model thereof, and hence must have a critical point κ that is measurable. More precisely:

Theorem 14. [Hamkins and Seabold, 2012] If U is an ultrafilter in V on the complete Boolean algebra $\mathbb{B}$, then the following are equivalent:

(1) $\check{V}_U$ is well-founded.

(2) $\check{V}_U$ is an ω-model (i.e. has the standard natural numbers).

(3) U meets all countable maximal antichains of $\mathbb{B}$ in V.

(4) U is countably complete over V (i.e. if $\langle a_n | n < \omega \rangle \in V$ is an ω-sequence of $a_n \in U$, then $\bigwedge_n a_n \in U$).

(5) U is weakly countably complete over V (i.e. if $\langle a_n | n < \omega \rangle \in V$ is an ω-sequence of $a_n \in U$, then $\bigwedge_n a_n \neq \varnothing$).

By (4), for the Boolean ultrapower to remain well-founded, we need significant large cardinal properties attaching to the completeness of the ultrafilter U.[56] Recall that many of the generic embeddings we wished to talk about had very small critical points. We are now in a position to mention the following:

Fact 15. Let U be an ultrafilter on a complete Boolean algebra $\mathbb{B} \in V$, and assume that $\check{V}_U$ is well-founded with j_U such that $crit(j_U) = \kappa$. Then $\check{V}_U$ cannot be used to interpret forcing constructions that change the structure of sets below V_κ.[57]

[56] The existence of a countably complete non-principal ultrafilter is equivalent to the existence of a measurable cardinal.

[57] This can be seen by the following quick:

Proof. It is a basic fact concerning measurable cardinals that if κ is measurable and the critical point of some j, then j preserves V_κ. (This is because if j_U did not, there would be (per impossibile) a different critical point below κ—see for example [Schindler, 2014], p. 51.) In the case where j_U is well-founded, since κ is the critical point of j_U we know that j_U preserves V_κ. Thus one cannot add subsets below V_κ. □

The key philosophical consequence of this fact is that if we are to interpret a forcing construction that involves the structure of sets below a measurable cardinal, then the Boolean ultrapower cannot be well-founded. This creates drawbacks with respect to the Naturalness Constraints; in virtue of our adherence to the iterative conception, one might think that well-foundedness is an especially important condition on structures conforming to our concept of set and qualifying as intended. But the above feature of the Boolean ultrapower violates this for many of our favourite forcing constructions. One cannot even add a single Cohen real without the Boolean ultrapower being non-well-founded. Moreover, we noted earlier that one of the key strengths of generic embeddings was to study 'small' accessible cardinals through the lens of large-cardinal-like embeddings. But, if we wish to move ω_1 using a generic embedding, then any Boolean ultrapower construction interpreting the required forcing will have differences in relatively 'small' levels' (e.g. $V_{\omega_1+2}^{\check{V}_U}$ and $V_{\omega_1+2}^{V^{\mathbb{B}}/U}$ will have different sets as $\omega_1^{\check{V}_U}$ is countable in $V^{\mathbb{B}}/U$). By the previous fact, there cannot be any difference in these levels without $\check{V}_U$ becoming non-well-founded; a philosophical limitation when we consider how a set theorist conceives of V and M as interrelated well-founded structures in $V[G]$. Philosophically, this feature of forcing is especially important—it keeps the models *well-founded* and the resulting model also therefore conforms to the iterative conception of set (a notion the Universist believes has *absolute* significance) in the sense that there is a powerset-like operation that can be iterated along the ordinals to generate the model.[58]

This philosophical problem has a technical manifestation: While we can still trace results derived from the study of these embeddings back through j_U to V, our reasoning in the embedded model may be restricted. Since the non-well-foundedness of the models implies a high degree of non-absoluteness (the satisfaction predicate itself is not even absolute[59]), we cannot simply use many of our normal assumptions regarding the relationship of sets in V to those in $\check{V}_U$ and $V^{\mathbb{B}}/U$. For example, consider the following reasoning template (stated in naive forcing-language):

Example 16. Extend V with a generic G that adds subsets above the least measurable κ to form $V[G]$. Now extend $V[G]$ with a generic H adding subsets below κ to form $V[G][H]$. Show that $V[G][H]$ satisfies some Δ_1-formula ϕ, and hence infer (by the absoluteness of Δ_1 formulas for transitive models) that V satisfies ϕ.

How would this be interpreted using the Boolean ultrapower? Letting $\mathbb{B}$ be the Boolean algebra for G, and U_G be the relevant ultrafilter, it is at least possible that '$V[G]$' can be interpreted as some well-founded inner model $V^{\mathbb{B}}/U_G$, extending $\check{V}_{U_G}$. Letting $\mathbb{C}$ be the Boolean algebra corresponding to H, $j_{U_G}(\mathbb{C})$ be the image of $\mathbb{C}$ under j_{U_G}, and U_H the corresponding ultrafilter on $j_{U_G}(\mathbb{C})$, we would then embed $V^{\mathbb{B}}/U_G$ into $(\widetilde{V^{\mathbb{B}}/U_G})_{U_H}$, with forcing extension $(V^{\mathbb{B}}/U_G)^{j_{U_G}(\mathbb{C})}/U_H$. But now, while we would know that $(V^{\mathbb{B}}/U_G)^{j_{U_G}(\mathbb{C})}/U_H \models \phi$, it would also be a non-well-founded structure, and so we cannot use the *absoluteness* and Δ_1 nature of ϕ to infer ϕ holds of $\check{V}_{U_G}$ (from where we could infer that ϕ holds of V either by absoluteness or pulling back along j_{U_G}). Rather we have to slowly and carefully trace where each embedding takes us, what is satisfied where, and the nature of the models involved. Specifically, since $(V^{\mathbb{B}}/U_G)^{j_{U_G}(\mathbb{C})}/U_H \models \phi$, we know by absoluteness there

[58] Plausibly, this is a reason why the independence phenomenon has gained more traction in set theory than in arithmetic; the arithmetic sentences we know of that are independent are require non-standard models to witness the independence.

[59] For details, see [Hamkins and Yang, 2013]. To give an example of just how extreme the phenomenon is, one can have two models that have *the same objects* as natural numbers, but disagree about whether a particular (non-standard) n is odd or even.

that $(\widetilde{V^{\mathbb{B}}/U_G})_{U_H} \models \phi$ (since the two models are transitive relative to each other). Pulling back along j_{U_H}, we then have that $V^{\mathbb{B}}/U_G \models \phi$, and hence by absoluteness $\check{V}_{U_G} \models \phi$. (We can then pull back again or just use absoluteness to obtain $V \models \phi$.) In short, if we repeatedly apply the Boolean ultrapower, our initial representation of V will be well-founded, as will our representation of $V[G]$. But then in the subsequent application of the Boolean ultrapower the new representation of $V[G]$ becomes ill-founded as does $V[G][H]$. The fact that our initial representation of V is a well-founded model, whereas our representation of $V[G][H]$ is not, means that we cannot infer by well-foundedness *directly* that our initial representation of V satisfies ϕ, instead we have to be careful where we apply absoluteness, making sure that we pull back along the Boolean ultrapower before applying absoluteness to ensure no illegitimate uses of the absoluteness of Δ_1-formulas. In the 'naive' reasoning, we just inferred from the fact that forcing keeps the models standard and $V[G][H] \models \phi$ that $V \models \phi$.

The philosophical point here is just the following; sometimes one might want to reason about V as though it were in a multiverse framework, but the moves made can depend upon the nature of forcing with respect to V. The fact that the Boolean ultrapower often fails to keep the models standard can result in extra steps needing to be made (rather than the single one required to interpret the forcing), and so can somewhat deform the set-theorist's intuitive thought.

One rejoinder to the above example is that there will be a representation of V inside the representation of $V[G][H]$, and one *can* conduct the absoluteness reasoning 'in one step' there before pulling back along the ultrapower (since *that* representation of V is ill-founded in the Universists real V, but transitive relative to the Boolean-ultrapower representation of $V[G][H]$). The issue is thus whether we should think of the set theorist as being located at a particular stage of a process in constructing a forcing argument, and try to keep her reasoning as natural as possible at every stage, or whether instead we should look for the best possible interpretation post hoc. If the former, then we encounter problems in absoluteness reasoning when interpreting the set theorist stepwise as she reasons, if the latter then this problem can be circumvented.

It seems then, that though the use of Boolean ultrapowers and quotient structures provides a way of modifying Boolean-valued structures into a two-valued framework, there can be a price, and this is high if one is especially attached to the well-foundedness of models in assessing the naturalness constraints. Interestingly, this opens a new area of enquiry; examine the cases where forcing and large cardinals combine to keep the ultrapower well-founded.[60] In these cases, the Boolean ultrapower fares *excellently* as an interpretation of forcing (since we can have well-founded, two-valued, proper-class-sized contexts of interpretation), suggesting that there are an especially 'natural' class of forcing constructions for the Universist. However, as a general technique for interpreting forcing constructions within a Universist framework, the technique may fall short.[61]

[60] See [Hamkins and Seabold, 2012] for discussion.

[61] A side remark should be made here about previous philosophical discussion concerning the Boolean ultrapower. Hamkins takes the Boolean ultrapower to show that the Universist can *simulate* forcing within her framework, and so should be moved to accept the existence of forcing extensions intuitively understood, and hence the falsity of her position. Koellner provides a response along the lines some of our observations here:

> "There are three important things to note about $[V^{\mathbb{B}}/U]$—it need not be transitive, it need not be well-founded, it is a definable class in V. For all three reasons it is as non-standard a model of set theory...one sees by construction that the model produced is not of the appropriate type to count as the universe of sets." ([Koellner, 2013], pp. 19–20)

3.4 Countable transitive models

Examining the Naturalness Constraints, we can see that given that (i) we actually want a model to be extended in a forcing-like way, (ii) these models should be two-valued, and (iii) we would like the relevant models to satisfy as much as possible of the theory of V, the strategy of interpreting forcing constructions via the use of a Boolean ultrapower got a lot right. It suffered two problems, however; one of scope (interpreting class-sized Boolean algebras) and one of how the ultrapower behaved with respect to 'small' sets. This is what we'll try to remedy by analysing how we might use countable transitive models in interpreting forcing. As we'll see, there are several respects in which the countable transitive model strategy performs better with respect to the Naturalness Constraints, but nonetheless this comes at a price (specifically concerning uncountable sets). We'll conclude that the countable transitive model strategy represents an attractive way of interpreting forcing using the symbol 'V', however for certain applications the Boolean-ultrapower strategy is especially natural.

The countable transitive model strategy comes in several forms. Initially, the method was designed to deal with the apparently problematic metamathematics of forcing from *within* **ZFC**. Given a desired relative consistency proof of some sentence ϕ, we assume that ϕ (respectively $\neg\phi$) is provable from **ZFC**. Since proofs are finite, we then know (by the Reflection, Löwenheim-Skolem, and Mostowski Collapse theorems) that there is a countable transitive model satisfying ϕ, and furthermore this countable transitive model can be arranged so as to have enough structure to enable forcing. Since the model is countable, we know that generics are available,[62] and can force to obtain $\neg\phi$ (whilst preserving the relevant fragment of **ZFC**),

Our point is just the following: Even if the Universist should (plainly) assert that neither $\check{V}_U$ nor $V^{\mathbb{B}}/U$ is the universe of sets, it can sometimes provide a very close simulation of forcing claims. Hamkins takes this to show that the Universist is wrong. We push this in the other direction: It shows that the Universist can sometimes capture much of the insight that is meant to be an advantage of the forcing multiversist's position.

[62] Assuming, of course, that V is uncountable.

contradicting the claim that (the relevant fragment of) **ZFC** proves ϕ.[63]

Such methods are fine as far as they go. For the purpose of allowing us to analyse forcing arguments establishing relative consistency proofs from within **ZFC** the technique performs adequately; any inconsistency of ϕ with **ZFC** could be traced to a countable model, and the relevant forcing argument shows that there is no such inconsistency (on the assumption that **ZFC** itself is consistent).

However, while V presumably satisfies **ZFC** and its finite fragmets for the Universist, there is no guarantee that a given countable transitive model of (a fragment of) **ZFC** represents V with respect to first-order truth *in general*. While we can, for any appropriate given finite set Γ of sentences satisfied by V, assume that there is a countable transitive model $\mathfrak{M}$ satisfying Γ, V can (and indeed will) satisfy sentences beyond those in Γ, and there may be sentences such that $V \models \phi$ and $\mathfrak{M} \models \neg\phi$. In order then to mimic the behaviour of V as closely as possible (and in line with the Naturalness Constraints) we would like a countable transitive model $\mathfrak{M}$ that resembles V sufficiently well to allow us to interpret forcing over V as closely as possible.

This can be achieved by assuming that there is a countable transitive model *elementarily equivalent* to V for first-order truth (we will denote such a model with the letter '$\mathfrak{V}$').[64] We can then interpret any forcing over this model, formulating axioms and proving theorems concerning $\mathfrak{V}$ (over which generics abound) and then, by the elementary equivalence, exporting the relevant consequences back to V proper. This is somewhat similar in spirit to the Boolean ultrapower; we find a context in which generics are available and subsequently pullback to V via a kind of resemblance. In the case of the Boolean ultrapower this was witnessed by an embedding, whereas in the case of our $\mathfrak{V}$ it is mere elementary equivalence (though see below for some remarks about embeddings).

For example, if we wish to formulate an axiom that uses an extension (say a generic embedding), we simply formulate it as concerned with $\mathfrak{V}$ (where extensions

[63]Formally, there are two main ways of executing the strategy (see [Kunen, 2013], IV.5.1 for details):

Suppose that we wish to prove that some statement ϕ is independent from **ZFC**. We then suppose that ϕ (or $\neg\phi$) has a proof in **ZFC** (from now on we only consider the case where we wish to show that $\neg\phi$ is unprovable). If $\neg\phi$ were provable, we would have $\mathbf{ZFC} + \phi \vdash 0 = 1$. Since proofs are finite, we then know that this proof would only use a finite set of axioms of **ZFC** (let it be denoted by 'Γ'). We then know (by the forcing method) that given such a Γ, there is a larger finite set of axioms of **ZFC** (let it be denoted by 'Λ') such that **ZFC** proves the conditional:

> "If there is countable transitive model for Λ, then there is a countable transitive model for $\Gamma + \phi$."

However, now we can use the Reflection Theorem, Löwenheim-Skolem Theorem, and Mostowski Collapse Lemma to then obtain a countable transitive model $\mathfrak{N}$ for Λ, and hence have a countable transitive model for $\Gamma + \phi$, contradicting our supposition that $\Gamma + \phi \vdash 0 = 1$.

Alternatively, we could expand $\mathscr{L}_{\in}$ to $\mathscr{L}_{\in,C,F}$ by two constant symbols C and F. We then add axioms to **ZFC** as follows (for this specific approach, see [Shoenfield, 1967] as well as [Kunen, 2013]):

Definition 17. **ZFC*** is a system in $\mathscr{L}_{\in,C,F}$ with the following axioms:

1. **ZFC**
2. C is a transitive set.
3. F is a bijection from ω onto C.
4. ϕ^C for every axiom ϕ of **ZFC** (note that, by Gödel's Second Incompleteness Theorem, this is an axiom scheme).

We know (by the Reflection Theorem, Löwenheim-Skolem Theorem, and Mostowski Collapse Lemma) that **ZFC*** is a conservative extension of **ZFC**. We can then treat C as our countable transitive model, and conduct our construction there.

[64]Uses of this idea for developing axioms have been explored by Friedman and collaborators in [Arrigoni and Friedman, 2013], [Antos et al., 2015], and [Friedman, 2016].

are uncontroversially available) and then know that any first-order consequence of the axiom true in $\mathfrak{V}$ is also true of V. If an ideal of the required kind I exists in V, then we will have a corresponding object $I' \in \mathfrak{V}$, and then there *will* be the required embedding $i' : \mathfrak{V} \longrightarrow \mathfrak{M} \subseteq \mathfrak{V}[G]$.

This strategy seems to resolve some of the problems of our previously considered accounts. First, regarding class forcing, we remarked earlier that we would like to interpret class forcings that lack a definable forcing relation, do not necessarily satisfy **ZFC**, or add classes but not sets. All these can be easily interpreted for $\mathfrak{V}$: Letting $\mathfrak{V} = (\mathbf{V}, \in)$, we can consider models of the form $\mathfrak{V}' = (\mathbf{V}, \in, \mathcal{C}^{\mathfrak{V}})$, where $\mathcal{C}^{\mathfrak{V}}$ is some countable collection of classes for $\mathfrak{V}$. We can then perfectly well consider extensions that add classes to $\mathcal{C}^{\mathfrak{V}}$ without adding sets to $\mathfrak{V}$, and/or ones in which the the forcing relation is not definable over $\mathfrak{V}$, and first-order consequences[65] thereby derived in $\mathfrak{V}$ can be transferred back to V. Similarly, since $\mathfrak{V}$ is very small by the standards of V, there is no obstacle to defining meaningful notions of Boolean completions for class forcings over $\mathfrak{V}$ (though they might not be in $\mathfrak{V}$), since we can use the resources in V external to $\mathfrak{V}$. Moreover, since generics are freely available over $\mathfrak{V}$ one does not *need* the Boolean algebras to interpret the forcing (though their study might nonetheless carry mathematical interest).

It seems then that the countable transitive model strategy is at least flexible concerning interpretation of forcing, providing a uniform way of interpreting the constructions. But how does it fare with respect to the Naturalness Constraints? Here, it also has a number of pleasing philosophical features. $\mathfrak{V}$ in many ways very closely resembles V; it is a two-valued structure that satisfies the same first-order sentences as V and (in contrast with certain applications of the Boolean ultrapower) is really a well-founded structure (and so forms a genuine cumulative hierarchy). In this way, $\mathfrak{V}$ provides a miniature 'picture' of V inside V, one which can be freely reasoned about using extensions. Moreover, once we have interpreted 'V' as '$\mathfrak{V}$' in a forcing construction, interpretation is completely transparent; $\mathfrak{V}$ *really is* extended by some G to form $\mathfrak{V}[G]$, and $\mathfrak{V}[G]$ has all the properties we would expect it to have: It is a well-founded cumulative hierarchy obtained from $\mathfrak{V}$ by forcing with G.

There are some drawbacks concerning the Naturalness Constraints, however. Unlike the Boolean ultrapower, $\mathfrak{V}$ cannot contain all the ordinals or uncountable sets. One response would be to argue that preservation of well-foundedness is preferable to containing uncountable sets. This is in part for philosophical reasons; a well-founded model is one that is obtained by admitting a genuine powerset-like operation transfinitely, and thus comes close to the iterative conception in an *absolute* sense. A non-well-founded model, even one containing uncountable sets, only has V_α *internally* (since the model will still satisfy the theorem of **ZFC** that every set belongs to some V_α), but actually admits of a descending membership sequence and so cannot be said to satisfying the iterative conception in any absolute sense. One might then think that having an absolute notion of the iterative conception should be privileged over the existence of actual uncountable sets—the former concerns the fundamentals of our contemporary conception of set whereas the latter is a *result* of our conception of set (as evidence for this, consider the fact that we call the Axiom of Foundation an *axiom*, whereas Cantor's Theorem is called a *theorem*).[66] We might thus feel that it is more important for the former to be satisfied in an absolute manner than the latter (it is more important that our model *actually* be a well-founded cumulative hierarchy

[65] In this paper we restrict only to *first-order* consequences. Salient here is that the use of countable transitive models can also be applied to higher-order axioms that use extensions (such as those mentioned in [Friedman, 2016]). See [Antos et al., S] for an application of these axioms to the Universist perspective.

[66] [Koellner, 2014], §1.4.1 makes a similar point concerning a different theorem (The Hydra Theorem) and axiom (that ϵ_0 is well-ordered).

than *actually* contain uncountable sets). Moreover, a well-founded model (even a countable one) can have (a subclass of) real membership as its membership relation, whereas a non-well-founded model cannot. Other reasons are more technical; as we noted earlier (when considering repeated applications of forcing to form $V[G][H]$, absoluteness inferences concerning Δ_1-formulas, and the Boolean ultrapower) one of the salient features of forcing is that it keeps the relevant model standard relative to the ground model. In the context of the use of countable transitive models then, repeated applications of forcing will always keep the model well-behaved with respect to the iterative conception (and subsequent uses of absoluteness), whereas this is not the case when the possibility of non-well-foundedness is introduced (as we saw with the abstract example of $V[G][H]$ earlier).

One question that then remains is the following: Is there a way of modifying the countable transitive model strategy so that we are able to incorporate *actually* uncountable sets? The following result shows that if countable transitive models are obtained in a certain way, then we can use them to construct uncountable contexts for forcing, under the assumption of large cardinals:

Theorem 18. [67] Assume that there is a proper class of measurable cardinals and that we have a $V_\delta \prec V$.[68] Let $\mathfrak{M}$ be a countable elementary submodel of V_δ, $\mathbb{P} \in \mathfrak{M}$ be a set forcing partial order such that $p \Vdash_{\mathbb{P}} \phi$ in $\mathfrak{M}$. Let π be the collapsing function to $\mathfrak{V} \equiv \mathfrak{M} \prec V_\delta \prec V$. Then there is a model $\mathfrak{V}^*$ such that:

(i) $\mathfrak{V}^*$ contains uncountable sets.

(ii) $\mathfrak{V}^*$ is transitive.

(iii) $\mathfrak{V}^*$ is elementarily equivalent to V.

(iv) There is a $G \in V$ such that G is $\mathfrak{V}^*$-generic for $\pi(\mathbb{P})$ and $\mathfrak{V}^*[G] \models \phi$.

As a rough sketch: The theorem can be shown by moving to a countable transitive model $\mathfrak{V}$, obtained by collapsing a countable elementary submodel of the universe and then iterating measures within $\mathfrak{V}$ uncountably many times.[69] While stated some-

[67] I am very grateful to Sy Friedman for suggesting the theorem and proof, and to Sandra Müller and Sam Roberts for some helpful clarifications.

[68] A $V_\delta \prec V$ can either be argued for informally or through the use of a truth predicate. See below for discussion.

[69] The details:

Proof. Assume that there is a proper class of measurable cardinals and that there is a $V_\delta \prec V$. Let $\mathfrak{M} \prec V_\delta \prec V$ be a countable elementary submodel, and π be the collapsing function to a countable transitive submodel $\mathfrak{V} \equiv V$. Take some forcing partial order $\mathbb{P} \in \mathfrak{M}$, and assume that there is a condition $p \in \mathbb{P}$ such that $p \Vdash_{\mathbb{P}} \phi$ over $\mathfrak{M}$. By the elementarity of π, there is a $q \in \pi(\mathbb{P})$ such that $q \Vdash_{\pi(\mathbb{P})} \phi$ over $\mathfrak{V}$. Since "There is a proper class of measurable cardinals" is first-order definable, $\mathfrak{V} \models$ "There is a proper class of measurable cardinals". So pick some $\mathfrak{V}$-measurable cardinal $\kappa_{\mathfrak{V}}$ such that $\kappa_{\mathfrak{V}} > rank^{\mathfrak{V}}(\pi(\mathbb{P}))$, and let $U_{\mathfrak{V}}$ be the measure witnessing $\kappa_{\mathfrak{V}}$'s $\mathfrak{V}$-measurability. Let U_V be $\pi^{-1}(U_{\mathfrak{V}})$, which is a measure on some measurable cardinal $\kappa_V \in V$. Since we know (by elementarity) that U_V is a genuine measure, we know that all iterated ultrapowers of V by U_V are well-founded. Moreover, it is a standard fact that any iteration of $\mathfrak{V}$ by $U_{\mathfrak{V}}$ embeds into an iteration of V by U_V (see here [Schindler, 2014], Ch. 10, especially Theorem 10.3 and Lemma 10.4), and so all successive ultrapowers of $\mathfrak{V}$ by $U_{\mathfrak{V}}$ are well-founded. So, iterate $\mathfrak{V}$ by $U_{\mathfrak{V}}$ uncountably many times to obtain $\mathfrak{V}^*$. Clearly $\mathfrak{V}^*$ is transitive and contains uncountable sets (since $U_{\mathfrak{V}}$ is iterable and we have embedded uncountably many critical points), and by the elementarity of the ultrapowers is elementarily equivalent to V. It just remains to show that there is a generic $G \in V$ that is $\mathfrak{V}^*$-generic for $\pi(\mathbb{P})$. Since $\mathfrak{V}$ was countable, we know that there is a $G \in V$ that is $\mathfrak{V}$-generic for $\pi(\mathbb{P})$. Moreover, it is a standard fact about well-founded ultrapowers that if a cardinal θ with measure U_θ is measurable in a transitive $\mathfrak{N}$, then $(V_\theta)^{\mathfrak{N}} = (V_\theta)^{Ult(\mathfrak{N}, U_\theta)}$. Thus $(V_{\kappa_{\mathfrak{V}}})^{\mathfrak{V}}$ is preserved at every stage of the iteration (and indeed in the direct limit), and so $\mathbb{P}$, the subsets of $\mathbb{P}$ (including any generics and dense sets) are preserved at each stage of the iteration. Thus G is generic for $\mathfrak{V}^*$. □

what technically, it shows that if we have some partial order $\mathbb{P}_V$ (definable using real parameters[70]) and $p_V \in \mathbb{P}_V$ such that $p_V \Vdash_{\mathbb{P}_V} \phi$, we can use countable transitive models to obtain an uncountable transitive model $\mathfrak{V}^*$ with a corresponding partial order $\mathbb{P}_{\mathfrak{V}^*}$, with a $p_{\mathfrak{V}^*} \in \mathbb{P}_{\mathfrak{V}^*}$ such that $p_{\mathfrak{V}^*} \Vdash_{\mathfrak{P}_{\mathfrak{V}^*}} \phi$, and *actual* $\mathfrak{V}^*$-generic G with $\mathfrak{V}^*[G] \models \phi$. In other words, given sufficient large cardinals, for any definable (with real parameters) set-forcing, we can always find an uncountable model elementarily equivalent to V and over which we can *actually* force with a very similar partial order. Thus, for a broad class of forcings, countable transitive models can also be instrumental in finding *uncountable* contexts for forcing, satisfying the Naturalness Constraints further.

Countable transitive models and Universism

Thus far, we have argued that a modification of the countable transitive model strategy is reasonably philosophically virtuous, at least as far as the Naturalness Constraints are concerned. We are not done yet, however. Critics of Universism (such as [Hamkins, 2012]) are well aware of the possibility of both a countable transitive model of **ZFC**, and well as one elementarily equivalent to V.[71] Hamkins has two main objections against the use of countable transitive models (especially with regards to forcing constructions).

First, he complains that the countable transitive model strategy is not broad enough, since it only accounts for forcing over certain models of set theory, whereas other methods (such as the Boolean ultrapower) can interpret (set) forcing over any model.[72]

Hamkins' second complaint surrounds the metamathematical commitments of the theorist endorsing the countable transitive model strategy. In order to interpret forcing over a countable transitive model, Hamkins argues, we have to assume the existence of a countable transitive model of **ZFC**. But this takes us somewhat beyond the consistency strength of our canonical set theory (i.e. **ZFC**); in fact the existence of a countable transitive model of **ZFC** implies the consistency of theories obtained by transfinite iterations of the consistency operator (so $Con(\mathbf{ZFC})$, $Con(Con(\mathbf{ZFC}))$, and so on transfinitely). In this way, he argues, the Universist is forced to pay a metamathematical tax in implementing the countable transitive model strategy.[73]

[70] Note that in the previous proof we can choose any finite list of real parameters to be in $\mathfrak{M}$.

[71] For the former claim see [Hamkins, 2012], and for the latter claim [Hamkins, 2003].

[72] See for example, the following remarks:

> "There are a number of drawbacks, however, to the countable transitive ground model approach to forcing. The first drawback is that it provides an understanding of forcing over only some models of set theory, whereas other accounts of forcing allow one to make sense of forcing over any model of set theory. With the countable transitive model approach to forcing, for example, the question "Is ϕ forceable?" appears sensible only when asked in connection with a countable transitive model M, and this is an impoverishment of the method." ([Hamkins, 2012], p421)

[73] For example, Hamkins argues as follows:

> "A second drawback concerns metamathematical issues surrounding the existence of countable transitive models of **ZFC**: the basic problem is that we cannot prove that there are any such models, because by Gödel's Incompleteness Theorem, if **ZFC** is consistent then it cannot prove that there are any models of **ZFC** at all. Even if we were to assume $Con(\mathbf{ZFC})$, then we still can't prove that there is a transitive model of **ZFC**, since the existence of such a model implies $Con(\mathbf{ZFC} + Con(\mathbf{ZFC}))$, and the consistency of this, and so on transfinitely...As a result, this approach to forcing seems to require one to pay a sort of tax just to implement the forcing method, starting with a stronger hypothesis than one

We find each of Hamkins' objections unconvincing from the Universist's perspective. Concerning the scope of the countable transitive model strategy: Whilst Hamkins is, as a matter of mathematical fact, correct that if we interpret width-extending constructions as concerned with countable transitive models, our interpretation of extensions will not apply to arbitrary models, in the present context this does not motivate the claim that insight might be missed. In particular, if we slowly work through the dialectic, we see that the main difficulties we noted were (i) to provide an interpretation of of axioms mentioning extensions, (ii) to explain why theorems making use of extensions were acceptable for proving facts about V, and do so (iii) in a way that accounts for the majority of our intuitive thinking concerning extensions. For *these* purposes, the countable transitive model strategy is reasonably satisfactory as it stands. It is uncontroversial for a Universist that some models cannot *literally* be extended in width, indeed it is *part* of their view that width extension is impossible for some models (such as V or one of its initial segments). Whilst it is true that certain constructions (such as use of the forcing relation or Boolean ultrapower map) permit interpretation of forcing extensions where 'V' denotes V, some uncountable class, or possibly even a non-well-founded structure, we deny that this represents a significant impoverishment of the countable transitive model strategy. Those techniques are always available in cases where the countable transitive model does not apply (such as in the case of a non-well-founded proper class structure). We do not wish to argue that the countable transitive model strategy is superior in *every* case, rather we wish to analyse what is *possible* given the Naturalness Constraints and argue that the countable transitive model strategy *is* natural for interpreting a *certain class* of extensions (namely when we wish to talk about V using well-founded and transitive extensions). Moreover, as we have seen, countable transitive models (if obtained by a collapse from a countable elementary submodel) can be *used* to obtain uncountable transitive contexts for set forcing. Our point is just that if we are concerned with V, the countable transitive model strategy provides ways of interpreting the forcing practice in line with the Naturalness Constraints.

One might instead interpret Hamkins as stressing not an impoverishment of the countable transitive model method itself, but rather the method of *forcing* as interpreted through countable transitive models.[74] One could not, for instance, then ask about the modal logic of forcing of V, since V is not a countable transitive model. Again, we wish to emphasise that we still *allow* other interpretations to be used. For example, in assessing questions concerning modal logic and forcing, we could instead talk about forcing relations, which will have to be mirrored between $\mathfrak{V}$ and V by elementary equivalence (it is just that $\mathfrak{V}$ has the actual extensions to go along with this). Our point is not what the Universist *must* do in interpreting forcing, but what she *can* do, and what the relative merits and drawbacks of each are.

These points do nothing to assuage Hamkins' second worry, however. There is still the charge that the Universist pays a problematic metamathematical tax. Of course, one blunt response would be to say that the tax is worth it, so it should be paid. We wish to see if we can do better, however.

There are at least two ways in which we might take the Universist to have to pay a tax she would rather not. The first is a *formal* one: it might be the case that she must use assumptions of greater strength than **ZFC** (or the theory in which one begins) in executing the strategy. Second, she might be paying an *ontological* tax, in that she has to commit herself to the existence of objects to which she would rather not.

ends up with just in order to carry out the argument." ([Hamkins, 2012], p421)

[74]Many thanks to an anonymous reviewer for suggesting this interpretation of Hamkins' criticisms.

Regarding the formal tax, Hamkins is correct that baldly asserting the existence of a countable transitive model of **ZFC** or one's starting theory (of which presumably $\mathfrak{V}$ is one) implies $Con(\mathbf{ZFC})$, $Con(Con(\mathbf{ZFC}))$, and so on, transfinitely. A response here is quick, however. The extent to which we move beyond **ZFC** is very dependent upon how we formulate the axiom. If we write:

Axiom 19. There exists a countable transitive model $\mathfrak{V}$ such that $\mathfrak{V} \models \mathbf{ZFC}$.

we would indeed go beyond **ZFC** (and, indeed V might satisfy very different sentences from $\mathfrak{V}$). Can we do better, getting a countable transitive model close in truth to to V, whilst not increasing the consistency strength of our theory? One has to be a little careful here; adding a constraint that $\mathfrak{V}$ satisfies *all* the same sentences (including parameters) as V would yield a contradiction.[75] We could instead formulate the axiom scheme that there is a V_α which is Σ_ω-elementary in V;[76] a very slight extension of the provable Lévy Reflection Principle, and then Skolemise over V_α. Much like the assumption of a transitive model of **ZFC** this assumption is very weak (well below an inaccessible). However, we can do even better formally. We can formulate the claim that there is a countable transitive model elementarily equivalent to V via the use of the following axiom *scheme* (as noticed by [Feferman, 1969]):

Definition 20. Let $\mathscr{L}_{\in,\bar{\mathfrak{V}}}$ be the language $\mathscr{L}_\in$ augmented with a single constant symbol $\bar{\mathfrak{V}}$. $\mathbf{ZFC}^{\bar{\mathfrak{V}}}$ is then a theory in $\mathscr{L}_{\in,\bar{\mathfrak{V}}}$ with the following axioms:

(i) **ZFC**

(ii) $\bar{\mathfrak{V}}$ is countable and transitive.

(iii) For every ϕ in $\mathscr{L}_\in$, $\phi \leftrightarrow \phi^{\bar{\mathfrak{V}}}$ (by Tarski's Theorem, this is an axiom scheme).

By replicating the proof of the Reflection Theorem for $\mathbf{ZFC}^{\bar{\mathfrak{V}}}$, we know that $\mathbf{ZFC}^{\mathfrak{V}}$ is a conservative extension of **ZFC**.[77,78] However, $\mathfrak{V}$ then satisfies *exactly* the same parameter-free first-order sentences of **ZFC** as V. When conducting a forcing proof in **ZFC** (or some extension thereof), we can use the usual trick of reflecting a finite set of axioms, but without loss of generality assume that they are reflected to $\mathfrak{V}$ to ensure enough similarity to V and conformity with the Naturalness Constraints.

[75] One way to see this is to note that $\omega_1^{\mathfrak{V}}$ would be uncountable according to $\mathfrak{V}$ but countable in V.

[76] Such an α is sometimes called "fully correct".

[77] *Proof Sketch.* Begin by replicating the usual proof of the Reflection Theorem for $\mathbf{ZFC}^{\bar{\mathfrak{V}}}$. Suppose then that $\phi \in \mathscr{L}_\in$ and $\mathbf{ZFC}^{\bar{\mathfrak{V}}} \vdash \phi$. Then there is a finite set of axioms Γ of $\mathbf{ZFC}^{\bar{\mathfrak{V}}}$, such that $\Gamma \vdash \phi$. Let ψ be the conjunction of all axioms of $\mathbf{ZFC}^{\bar{\mathfrak{V}}}$ in Γ containing $\bar{\mathfrak{V}}$, and χ be conjunction of all axioms of $\mathbf{ZFC}^{\bar{\mathfrak{V}}}$ in Γ not containing $\bar{\mathfrak{V}}$. Without loss of generality, we may then assume that $\Gamma = \{\psi, \chi\}$. ψ is then a sentence which states that there is a countable transitive model of some finite list of **ZFC** axioms Λ, and χ is a finite part of **ZFC**. Since **ZFC** already proves that there is a countable transitive model for Λ (by the Reflection Theorem and Mostowski Collapse Lemma in **ZFC**), and hence proves $\chi \wedge (\exists \mathfrak{M})\mathfrak{M} \models \Lambda$, we have $\mathbf{ZFC} \vdash \phi$. Many thanks to Asaf Karagila and Sam Roberts for pointing out to me this theorem of Feferman, and for discussion of the proof. A referee helpfully points out that it is possible that the models of $\mathbf{ZFC}^{\bar{\mathfrak{V}}}$ involved may be non-well-founded, and in fact that it may be that the *all* the models of $\mathbf{ZFC}^{\bar{\mathfrak{V}}}$ in a universe of set theory are non-well-founded. We wish to emphasise that this proof sketch should *only* convince the Universist of the conservativity of the theory, it is *not* intended to convince her that countable transitive models of the required kind actually exist. That claim will require other philosophical and mathematical arguments (such as those given in the text); the point here is that we can *formally* get away with no increase in consistency strength, if we think that a countable transitive model elementarily equivalent to V exists.

[78] It is interesting to note that [Hamkins, 2003] discusses this formulation in depth, and is well aware of the conservativity result. It is perhaps the metatheoretical character of the claim that $\mathfrak{V}$ satisfies **ZFC** that explains the absence of this strategy when considering countable transitive models in [Hamkins, 2012].

Whilst we can only see from the metatheory (without a non-conservative extension of **ZFC**) that $\mathfrak{V}$ is a countable transitive model elementarily equivalent the V, this is to be expected: V cannot have access to its own truth definition. Moreover, the Universist is no stranger to accepting the use of metatheoretical talk in her set-theoretic reasoning (for example: "$V \models \mathbf{ZFC}$").

So it seems that we need not pay a significant *formal* price when interpreting width extensions with countable transitive models; the formal price can be circumvented by adopting a particular formulation of the countable transitive model strategy. It remains to be seen that the Universist need not pay an *ontological* price. To what extent does the existence of a countable transitive model elementarily equivalent to V represent a substantial ontological presupposition for the Universist? We shall argue that the Universist *already* has good reasons to accept the existence of such a model given her commitment to the existence of a unique and maximal proper class model of set theory. We shall pursue two strategies here: (1) examining Skolem functions in the Universist's metatheory, and (2) the use of a truth predicate.

For (1) we start by examining an informal argument of Cohen[79] for the existence of a countable transitive model of **ZF**:

> "The Löwenheim-Skolem theorem allows us to pass to countable submodels of a given model. Now, the "universe" does not form a set and so we cannot, in **ZF**, prove the existence of a countable sub-model. However, informally we can repeat the proof of the theorem. We recall that the proof merely consisted of choosing successively sets which satisfied certain properties, if such a set existed. In **ZF** we can do this process finitely often. There is no reason to believe that in the real world this process cannot be done countably many times and thus yield a countable standard model for **ZF**." ([Cohen, 1966], p79)

While Cohen is primarily interested in the existence of a countable transitive model for **ZF**, we can import his argument to the case of V as follows. When considering finite subsets Γ of sentences in $\mathscr{L}_{\in}$ satisfied by V we move to a countable transitive model of Γ by using the Reflection Theorem to find a $V_\alpha \models \Gamma$. We then use AC to find a set of functions F^{V_α} such that for every existential statement $\exists x \phi(x)$ true in V_α, there is an $f_\phi \in F$ that picks a single witness a such that $\phi(a)$ holds. We then form a countable $\mathfrak{M} \models \Gamma$. Since the relation (i.e. $\in$) on V_α is extensional, well-founded, and set-like, so is the relation on $\mathfrak{M}$, and we then use the Mostowski Collapse Lemma to collapse to obtain a countable transitive model $\mathfrak{M}' \models \Gamma$. Turning now to V, we simply note that V is one model of **ZFC** among many. Thus, we can posit the existence of a class of Skolem functions F^V for V (by 'choosing[80]' a witness for the countably many existential statements satisfied by V with a countable sequence of choices). Then, by Skolemising and Collapsing, we obtain a countable transitive model $\mathfrak{V}$ that satisfies exactly the same parameter-free first-order sentences as V. The main issue here is that, by Tarski's Theorem on the undefinability of truth, that V does not know that F^V provides its own Skolem-functions. Thus, by adopting this *informal* and *metatheoretic* version of the Löwenheim-Skolem Theorem, we justify the existence of a countable transitive model for V. Again, we note that the Universist should not be bothered by making these various metatheoretic claims in virtue of the things she wants to say *already* concerning her own view.

[79] Similar arguments (from a non-Universist perspective) are available in [Arrigoni and Friedman, 2013].

[80] As Cohen notes, this is not possible in $\mathbf{ZF}(\mathbf{C})$ by Gödel's Second Incompleteness Theorem.

Though the Cohen-style argument depends on only mild metatheoretic assumptions (the possibility of transferring a countable sequence of choices in the metalanguage to the existence of Skolem functions in V), it nonetheless depends upon a certain amount of combinatorial reasoning outside the object language. We might, therefore, find a more *logical* approach desirable. Here we find a role for the use of a *truth predicate*. We begin by noting that the Universist already makes claims concerning first-order *truth* in V. For example:

Example 21. "In V every sentence is either true or false."

Example 22. "Every axiom of **ZFC** is true in V."

Given that she would like some formal apparatus for talking about her own view, it is natural for her to countenance the existence of a truth predicate for talking about her own view (otherwise she ends up baldly stating the sentences in question, rather than the truth thereof). We therefore make the following:[81]

Definition 23. Let $\mathbf{ZFC}_{Tr}$ be the result of adding to **ZFC** a truth-predicate $Tr(x)$ for a satisfaction class Tr for first-order truth, such that:

(i) $Tr(x)$ iff $x \in Tr$.

(ii) Tr is a class consisting of pairs $\langle \ulcorner \phi \urcorner, \vec{a} \rangle$, where $\vec{a}$ is an assignment of free variables in ϕ to the objects ordered by $\vec{a}$.

(iii) Tr is correct about the atomic truths of V in the language of **ZFC**.

(iv) $\langle \ulcorner \neg \phi \urcorner, \vec{a} \rangle \in Tr$ iff $\langle \ulcorner \phi \urcorner, \vec{a} \rangle \notin Tr$.

(v) $\langle \ulcorner \phi \wedge \psi \urcorner, \vec{a} \rangle \in Tr$ iff $\langle \ulcorner \phi \urcorner, \vec{a} \rangle \in Tr$ and $\langle \ulcorner \psi \urcorner, \vec{a} \rangle \in Tr$.

(vi) $\langle \ulcorner \exists x \phi(x) \urcorner, \vec{a} \rangle \in Tr$ iff for some b, $\langle \ulcorner \phi \urcorner, b^\frown \vec{a} \rangle \in Tr$.

(vii) We allow the use of $Tr(x)$ in the Comprehension and Replacement Schema.

In other words $Tr(x)$ is a predicate corresponding to a class computing atomic truth correctly and closed under the Tarskian recursive clauses.

$\mathbf{ZFC}_{Tr}$ represents a very minimal formalisation of the first-level of typed truth over V for the Universist. We now mention the following:

Fact 24. $\mathbf{ZFC}_{Tr}$ proves that there is a countable transitive model elementarily equivalent to V for first-order parameter-free truth in **ZFC**.[82]

Thus, if we permit the addition of a truth predicate for V to **ZFC**, the use of which a Universist may well accept (given the other presuppositions of her view), we obtain a countable transitive model elementarily equivalent to V. This argument, in combination with the informal use of the Löwenheim-Skolem Theorem, shows that the existence of a countable transitive model elementarily equivalent to V is little

[81] I am grateful to Sam Roberts for technical discussion concerning the use of truth predicates and countable transitive models.

[82] This can be shown as follows:

Proof. (Sketch) Begin by proving the Reflection Theorem for $\mathbf{ZFC}_{Tr}$. Then, reflect on the formula "$\ulcorner \phi \urcorner$ is a Gödel code of a formula and $Tr(\ulcorner \phi \urcorner)$" to obtain a V_α elementarily equivalent to V (since for a given natural number $n = \ulcorner \phi \urcorner$, $V \models Tr(\ulcorner \phi \urcorner)$ iff $V_\alpha \models Tr(\ulcorner \phi \urcorner)$, and hence $V \models \phi$ iff $V_\alpha \models \phi$). Then Skolemise and Collapse over V_α to get the desired countable transtive model. □

ontological price to pay, if any. While the Universist still pays a 'tax' in the sense that the theory she forces over (i.e. the **ZFC**-theory of V) is different from the one used to construct the countable transitive model (i.e. $\mathbf{ZFC}_{Tr}$), the tax is both one she is already happy to pay and enables her to force over her preferred theory of sets.

Conclusion

We have seen that Universist has some reason to want to use forcing to prove theorems and formulate axioms about V. Whilst the use of a Boolean ultrapower sometimes results in non-well-founded models, it suggests a class of forcings of special interest for the Universist; those for which the Boolean ultrapower can be kept well-founded, and thus she has a particularly natural interpretation of forcing. The countable transitive model strategy, augmented to yield a model elementarily equivalent to V, we found to be an effective, flexible, and well-motivated method for interpreting forcing constructions and can be modified to yield uncountable contexts over which one can force. The extent to which one thinks that Hamkins' complaints have dialectic force against the Universist, thus turns precisely on further constraints on naturalness of interpretation. Perhaps such an argument can be made, and we do not expect this clarification of the tools available to the Universist in interpreting forcing to be convincing to the Forcing Multiversist; she will always regard her interpretation as the *most* natural, since the Universist will always have to admit *some* re-interpretation of language. For now, however, we take ourselves to have shown that the Universist should remain largely unfazed by criticisms stemming from forcing constructions.

References

[Antos, 2018] Antos, C. (2018). *Class Forcing in Class Theory*, pages 1–16. Springer International Publishing, Cham.

[Antos et al., S] Antos, C., Barton, N., and Friedman, S.-D. (S). Universism and extensions of V. Manuscript under review.

[Antos et al., 2015] Antos, C., Friedman, S.-D., Honzik, R., and Ternullo, C. (2015). Multiverse conceptions in set theory. *Synthese*, 192(8):2463–2488.

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

[Baumgartner and Hajnal, 1973] Baumgartner, J. and Hajnal, A. (1973). A proof (involving Martin's Axiom) of a partition relation. *Fundamenta Mathematicae*, 78(3):193–203.

[Cohen, 1966] Cohen, P. (1966). *Set Theory and The Continuum Hypothesis*. W.A. Benjamin, Inc.

[Cummings, 2010] Cummings, J. (2010). *Iterated Forcing and Elementary Embeddings*, pages 775–883. Springer Netherlands, Dordrecht.

[Feferman, 1969] Feferman, S. (1969). Set-theoretical foundations of category theory. In *Reports of the Midwest Category Seminar III, Lecture Notes in Mathematics Volume 106*, pages 201–247. Springer.

[Foreman, 1986] Foreman, M. (1986). Potent axioms. *Transactions of the American Mathematical Society*, 294(1):1–28.

[Foreman, 1998] Foreman, M. (1998). Generic large cardinals: New axioms for mathematics? *Documenta Mathematica*, 2:11–21.

[Foreman, 2010] Foreman, M. (2010). Ideals and generic elementary embeddings. In Kanamori, A. and Foreman, M., editors, *Handbook of Set Theory*, pages 885–1147. Springer.

[Friedman, 2000] Friedman, S.-D. (2000). *Fine Structure and Class Forcing*. de Gruyter. de Gruyter Series in Logic and its Applications, Vol. 3.

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

[Friedman, 2010] Friedman, S.-D. (2010). Constructibility and class forcing. In Kanamori, A., editor, *Handbook of Set Theory*, pages 557–604. Springer.

[Friedman, 2016] Friedman, S.-D. (2016). Evidence for set-theoretic truth and the hyperuniverse programme. *IfCoLog J. Log. Appl.*, 4("Proof, Truth, Computation" 3):517–555.

[Gitman et al., 2017] Gitman, V., Hamkins, J. D., Holy, P., Schlicht, P., and Williams, K. (2017). The exact strength of the class forcing theorem.

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

[Hamkins, 2003] Hamkins, J. D. (2003). A simple maximality principle. *The Journal of Symbolic Logic*, 68(2):527–550.

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

[Hamkins and Seabold, 2012] Hamkins, J. D. and Seabold, D. E. (2012). Well-founded Boolean ultrapowers as large cardinal embeddings. arXiv:1206.6075 [math.LO].

[Hamkins and Yang, 2013] Hamkins, J. D. and Yang, R. (2013). Satisfaction is not absolute. arXiv:1312.0670v1 [math.LO].

[Holy et al., 2016] Holy, P., Krapf, R., Lücke, P., Njegomir, A., and Schlicht, P. (2016). Class forcing, the Forcing Theorem, and Boolean completions. *The Journal of Symbolic Logic*, 81(4):1500–1530.

[Holy et al., 2018] Holy, P., Krapf, R., and Schlicht, P. (2018). Characterizations of pretameness and the ord-cc. *Annals of Pure and Applied Logic*, 169(8):775 – 802.

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

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

[Koellner, 2013] Koellner, P. (2013). Hamkins on the multiverse. In Koellner, P., editor, *Exploring the Frontiers of Incompleteness*.

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

[Kunen, 1971] Kunen, K. (1971). Elementary embeddings and infinitary combinatorics. *The Journal of Symbolic Logic*, 36:407–413.

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

[Larson, 2004] Larson, P. (2004). *The Stationary Tower: Notes on a Course by W. Hugh Woodin*. University lecture series. American Mathematical Society.

[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.

[Painlevé, 1900] Painlevé, P. (1900). *Analyse des travaux scientifiques*. Gauthier-Villars.

[Schindler, 2000] Schindler, R. (2000). Proper forcing and remarkable cardinals: I. *Bulletin of Symbolic Logic*, 6(2):176–184.

[Schindler, 2001] Schindler, R. (2001). Proper forcing and remarkable cardinals: II. *Journal of Symbolic Logc*, 66(3):1481–1492.

[Schindler, 2014] Schindler, R. (2014). *Set Theory: Exploring Independence and Truth*. Springer.

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

[Smullyan and Fitting, 1996] Smullyan, R. M. and Fitting, M. (1996). *Set Theory and the Continuum Problem*. Clarendon Press.

[Steel, 1996] Steel, J. (1996). The core model iterability problem. In *Lecture Notes in Logic*, volume 8. Springer.

[Todorčević and Farah, 1995] Todorčević, S. and Farah, I. (1995). *Some Applications of the Method of Forcing*. Yenisei, Moscow.