BICONTEXTUALISM FOR TRUTH AND SETS

An Intermediate Position between Generality Absolutism and Generality Relativism

SIMON SCHMITT
DEPARTMENTS OF PHILOSOPHY AND EDUCATION
CENTER FOR LOGIC, LANGUAGE, AND COGNITION

UNIVERSITÀ
DI TORINO

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Ph.D.)

Under the kind guidance of

Andrea Iacona & Lorenzo Rossi

Reviewed by

Francesca Boccuni & Øystein Linnebo

June 29, 2025

DEFENSE COMMITTEE

This thesis is presented before the following committee:

Francesca Boccuni

Øystein Linnebo

Lorenzo Rossi

The defense is scheduled to take place on
July 4, 2025
at the University of Turin, Italy.

ABSTRACT

Generality absolutism—the thesis that quantifiers can range over absolutely everything there is—has attracted significant philosophical attention in recent decades. There are several reasons for this. First, absolute generality has implications for the foundations of semantics, particularly for definitions of key concepts such as logical consequence. Second, the study of quantifiers' range and meaning plays a crucial role in the interpretation of natural language. Third, since quantifiers and formal logic are central to mathematics, generality absolutism also concerns the scope of mathematical assertions, which is particularly significant within the foundations of mathematics.

The main opponent of generality absolutism is generality relativism, which holds that quantifiers are always restricted. The strongest argument for relativism is based on paradox: if successful, this argument undermines absolutism and has consequences just as far-reaching as those that would follow from adopting absolutism itself.

The aim of this thesis is to develop an intermediate position—called *bicontextualism*—that is located between absolutism and relativism. The core idea of bicontextualism is that both positions capture important insights: relativism is justified insofar as quantifiers involved in paradoxes must be restricted, whereas absolutism is justified insofar as quantifiers used in sentences not related to paradoxes require no restriction and may range over absolutely everything there is.

Bicontextualism is motivated by the fact that relativistic solutions to the set-theoretic and truth-theoretic paradoxes yield valuable insights into the nature of the underlying concepts—insights that should be preserved within the intermediate position. The most important insight is that the concepts 'truth' and 'set' are *hierarchical*. Any satisfactory solution of the paradoxes must respect and maintain this hierarchical character.

The thesis is structured as follows:

Chapter 1 presents a more detailed introduction to the topic.

Chapter 2 lays out the philosophical background for this intermediate position. It introduces both, absolutism and relativism, and examines in detail the relativist argument from paradox. It is argued that the relativist conclusion—that *all* quantifiers must be restricted—is overly hasty and unwarranted. It also identifies the key insights from the relativistic solutions—the hierarchical nature of 'truth' and 'set'—that is to be preserved by the intermediate position.

Chapter 3 provides the formal background necessary for the development of bicontextualism. This includes foundational material on logic, type-free truth, set theory, and an alternative approach to model theory called the *RU approach*.

Chapters 4 and 5 spell out the bicontextualist position within the contexts of truth theory and set theory, respectively. Chapter 4 develops a type-free theory of truth for higher-order object languages that is compatible with generality absolutism and keeps the hierarchical nature of the concept of truth. Chap-

ter 5 presents a semantics for set theory that likewise aligns with generality absolutism and preserves the cumulative-iterative conception of 'set'.

Chapter 6 reflects on the philosophical implications of bicontextualism, focusing on the foundations of mathematics and the semantics of higher-order languages. In particular, it develops a first step towards a consistent interpretation of higher-order expressions that is compatible with generality absolutism. In addition, it compares bicontextualism to other approaches found in the literature. Finally, the chapter mentions potential directions for future research and open questions that arise from preceding chapters.

Chapter 7 concludes and reevaluates the goals set in chapters 1 – 5 and assesses to what extent they have been achieved.

ACKNOWLEDGMENTS

It is a pleasure to look back and think of everyone who supported me during the last years.

First and foremost, I am deeply grateful to my supervisor Lorenzo Rossi, who has always generously shared his ideas with me. But there is more than that: writing a PhD thesis can be challenging, and Lorenzo's support extended far beyond intellectual matters. Thank you, Lorenzo—you have helped me grow as a philosopher, a logician, and a researcher more generally. I am also very thankful to my supervisor Andrea Iacona for his invaluable feedback, suggestions, and thoughtful comments, which have greatly improved this work. Special thanks to Julien Murzi for many extensive discussions and critical insights, which greatly improved this work. Special thanks also to Hannes Leitgeb, who needed only seconds to spot weaknesses in early drafts and whose suggestions for better ways to present the material have greatly improved this thesis. I am especially grateful to the referees, Francesca Boccuni and Øystein Linnebo, for their extensive comments and many valuable suggestions that helped improve this thesis. I gratefully acknowledge the financial support of the Northwest Italy Philosophy PhD Program (FINO), whose scholarship made this work possible.

I have had the pleasure of working with and learning from many others while developing this project. I list them in alphabetical order and sincerely apologize to anyone I may have unintentionally left out: Neil Barton, Tim Button, Matteo de Ceglie, Ahmed Çevik, Laura Crosilla, Jann Paul Engler and the entire reading group on Generality and Impredicativity, Martin Fischer, O. Foisch, David Hofmann, Leon Horsten, Deborah Kant, Giorgio Lenta, Angelica Mezzadri, Matteo Plebani, Sam Roberts, Lucas Rosenblatt, Hosea von Hauff, Chris Scambler, Jan Sprenger, Davide Sutto, Giorgio Venturi, as well as the audiences in Oslo, London, Warsaw, Pavia, Munich, Konstanz, and Turin.

I've been fortunate to receive personal support outside of academia, and I'm deeply grateful to those who have been there for me throughout these years. To begin with, Jacqueline—thank you for your unwavering support. I am especially grateful for your patience in proofreading this thesis—despite your claim that it's the most boring thing you've ever read! Your presence and encouragement over the past years truly means the world to me. Next, Benny, your support was essential. I genuinely don't know how I would have managed without it. Thank you so much. And last, but definitely not least, my parents, Jeanette and Dieter, who have supported me countless times when things got tough. Thank you both so much.

CONTENTS

1

INTRODUCTION

Paradoxes arise in many areas, including various scientific disciplines. We also encounter them in everyday language—for example, the famous barber paradox: Who shaves the barber who shaves everyone who does not shave themselves? Most importantly, however, paradoxes appear in semantics and in the foundations of mathematics. Typically, they reveal that fundamental assumptions about certain seemingly innocent concepts may, upon closer examination, turn out to be far more problematic than initially assumed.

For instance, it appears perfectly natural to assume that for any sentence φ, the sentences φ and "φ' is true' should be interchangeable. Yet, as philosophers and logicians have discovered, this leads to the problematic *Liar paradox*: what about the sentence λ that asserts of itself that it is false?

From a naïve perspective, any sentence should be either true or false, and thus λ must be either true or false. But then, we can reason from "λ' is true' to λ because they are interchangeable, and then from λ to "λ' is false' by the definition of λ. Similarly, we can reason from "λ' is false' to λ by the definition of λ, and then to "λ' is true' because they are interchangeable. Thus, we are forced to conclude that λ is true if and only if it is false—an outright contradiction. Clearly, something has gone wrong.

A similar problem arises in the foundations of mathematics. It seems natural to assume that for any property $\psi(x)$, there exists a set containing all and only those objects for which ψ holds.[1] But what about the property of being non-self-membered? If there is a set of all non-self-membered objects, is it a member of itself? A moment's reflection reveals that we end up in a similar situation: it is a member of itself if and only if it is not. This is the famous *Russell paradox*. Again, something must have gone wrong.

As anticipated, the lesson to be learned is that seemingly innocent naïve principles turn out to be much more problematic then initially assumed: in the case of the Liar paradox, the unrestricted application of the truth predicate to self-referential sentences leads to contradictions, so the notion of truth must be carefully restricted. In the case of Russell's paradox, unrestricted set formation leads to contradictions, so the definition of sets must be properly constrained.

None of this is new; the literature on the Liar and Russell's paradoxes is vast, and since their discovery, philosophers, logicians, and mathematicians have developed a multitude of sophisticated approaches and solutions. Their motivation should be clear from the foregoing: the principles of truth and membership lie at the heart of semantics and mathematics. If these fundamental principles are not well-understood, how can we build more complex semantic and mathematical theories on top of them?

1 In what follows, I speak only about sets, but the reader should keep in mind that the considerations apply equally to any collection-like entity. In this sense, I use the term 'set' synonymously with 'collection'.

While it is clear to some extent how the paradoxes can be addressed—though there is no consensus regarding what the *right* solution is—it is less clear what exactly the implications of these paradoxes are. In particular, one prominent alleged implication of the paradoxes—as has been argued—is that *generality absolutism*, the claim that quantifiers can range over absolutely everything there is, must be false. Generality relativists argue that Russell's paradox implies that not all objects can be gathered into a single domain, and therefore quantification must always be restricted. I call this the *relativistic argument from paradox*.

Here is the argument in a nutshell; details will be provided in section 2.2.1: the relativistic argument from paradox holds that Russell's paradox implies the impossibility of an all-encompassing domain containing absolutely everything over which we can quantify without contradiction. Since the paradox arises when we attempt to consider 'the set of all sets that do not contain themselves', it reveals that assuming a fixed, absolute totality of all objects (which would have to contain all sets) leads to inconsistency. Therefore, quantification must always be understood as restricted to a non-absolute domain that cannot be the totality of all objects. Consequently, the relativists argue, quantification is inherently relative rather than absolute.

Yet it is important to ask whether this conclusion is genuinely justified. Indeed, one of the central questions of this work is precisely this:

> **Question.** Is the relativistic argument from paradox correct? That is, does Russell's paradox truly imply that generality absolutism must be false?

To give an immediate spoiler: the answer is *no*, the conclusion is not justified. Russell's paradox—or any other membership or collection paradox—does not imply that generality absolutism must be false.

Nevertheless, the most sophisticated solutions to the paradoxes are inherently relativistic. *Truth-theoretical contextualism* traces the seemingly simultaneous truth and falsity of the Liar sentence to a (possibly hidden) contextual element. On that approach, a contextualized version of the Liar sentence is not true and false in a single context but turns out false in one context and true in another: the contextualized Liar is rephrased as 'λ_0 does not say anything *true-in-*c_0.' Then, λ_0 is not true in c_0. However, contextualists claim, the evaluation of λ_0 induces a context shift to a subsequent context c_1 in which λ_0 turns out true instead.[2] The context shift involves a domain expansion. Moreover, any context can be extended in this way. Consequently, any domain can be expanded, and so no domain can achieve absolute generality.

In the set-theoretic case, according to the *cumulative-iterative conception of set*, sets are arranged stage-wise in an open-ended hierarchy of ever-growing ranks. Moreover, on that account, set comprehension is always restricted to previously given sets. Russell sets can be formed at any stage, but they enter the hierarchy (as a set) only at a later stage. This, however, implies that each rank is open for expansion, and no rank can achieve absolute generality.

At this point, it seems as if absolutism and the solutions to the paradoxes are in tension: if we adopt the sophisticated contextualist or cumulative-iterative

2 More on this in section 2.2.2.

approaches to the notions of truth and membership, we must give up on absolutism. Alternatively, if we adopt absolutism, we must give up on the sophisticated contextualist or cumulative-iterative approaches to the notions of truth and membership. But we cannot have both—or so it seems.

In a recent publication, Lorenzo Rossi [144] has shown that, in the truth-theoretical setting, this dilemma is by no means forced upon us, and that there is a way to combine the view that at least *some* sentences can be given an absolutely unrestricted interpretation with contextualist, broadly hierarchical solutions to the paradoxes. Following Richard Cartwright [21], Rossi drops the assumption that domains have to be entities (such as sets), and instead advocates an ontologically neutral reading of the notion 'domain'. Moreover, he proposes to divide the sentences of the underlying language into those that lead to paradoxes, such as the Liar sentence, and those that are unproblematic in this respect, such as, e.g., '$5 + 7 = 12$' or 'Everything is self-identical'. He then proposes a *bipartite* semantics, where unproblematic sentences are interpreted over an all-inclusive (non-objectual) 'domain', while problematic sentences are interpreted over restricted, set-sized domains.[3] Thus, as Rossi shows, we can combine absolutism for unproblematic sentences with contextualist approaches to the semantic paradoxes.

It is one of the central aims of this work to articulate and develop this view in detail. In chapter 2, I introduce three positions: *generality absolutism*, *generality relativism*, and *bicontextualism*—the intermediate position developed in Rossi [144]. I offer a detailed examination of the relativistic argument from paradox and, following Rossi, argue that a crucial assumption underlying this argument is that domains must be *things*, i.e., entities such as sets. Once this assumption is rejected, the relativistic argument from paradox can be effectively resisted. Furthermore, I analyze the Liar and Russell's paradoxes in depth and argue that both share a common root: the inherently hierarchical nature of the underlying concepts. I argue that this hierarchical character of 'truth' and 'set' is a fundamental lesson to be drawn from the paradoxes, and that any adequate position that is intermediate between absolutism and relativism must preserve this feature. Finally, I argue that Rossi's bicontextualism offers precisely such a middle ground: it combines the hierarchical solutions to the paradoxes with an unrestricted 'domain' of quantification, at least for *some* sentences.

Further aims of this work include developing this position within both a truth-theoretical and a set-theoretical context. To this end, chapter 3 introduces the necessary technical background. In the truth-theoretical setting, this includes the theory of inductive definitions underlying Kripke's construction, as well as a sophisticated form of contextualism due to Charles Parsons [118], Michael Glanzberg [47], and Lorenzo Rossi [144]. In the set-theoretical context, I review the distinction between intended and unintended models of first- and second-order set theory, with a special focus on Zermelo's quasi-categoricity theorem. Moreover, since developing bicontextualism as a semantic theory compatible with absolutism requires non-standard semantic tools, I introduce the *Rayo-*

3 In the following, 'domain' (with quotation marks) refers to non-objectual domains, while domain (without quotation marks) refers to objectual domains. This distinction will be crucial throughout this work and will be discussed in much greater detail in the upcoming chapters.

Uzquiano approach (RU approach): a higher-order, non-objectual alternative to standard model theory developed by Agustín Rayo and Gabriel Uzquiano [135]. The final section of chapter 3 develops the RU approach in detail.

With this groundwork laid, I proceed to develop bicontextualism for both truth and set theory. The case for truth will be addressed first. While Rossi has developed bicontextualism only for first-order object theories, I argue that there are compelling reasons to extend formal semantics to capture higher-order object languages. Only higher-order logic enables us to speak about mathematical structures in a way that ensures their uniqueness (up to isomorphism). In chapter 4, I extend truth-theoretic bicontextualism to higher-order languages: first, I motivate the need for higher-order languages in this context, then I provide a formal presentation of the framework, and finally, I address potential objections to incorporating higher-order languages into formal semantics.

The next task will be to develop a set-theoretic version of bicontextualism. Standard semantics for set theory is incompatible with unrestricted quantification over all sets. However, there are strong reasons to endorse such quantification: after all, the subject matter of set theory is *all sets*, and only with an unrestricted interpretation of the quantifiers can set theory have its intended interpretation. In chapter 5, I first motivate this position, then present a formal account of set-theoretic bicontextualism, and finally engage with possible objections to the view.

Once the technical development is complete, I turn to philosophical implications in chapter 6. This includes an exploration of an absolutist-friendly interpretation of higher-order expressions and further reflections on the specific consequences for truth-theoretic and set-theoretic contexts. Finally, I compare the position defended here with competing approaches in the literature. Chapter 7 will summarize and conclude the work.

2

GENERALITY, PARADOXES, BICONTEXTUALISM

Generality absolutism is the claim that it is possible to quantify over absolutely everything there is. *Generality relativism*, on the other hand, is the claim that quantification is always restricted to a less-than all-inclusive domain.

Bicontextualism is located in between absolutism and relativism, and it is the main aim of this chapter to introduce bicontextualism as a natural and compelling intermediate position. In a nutshell, bicontextualism is the position that only some sentences can achieve absolute generality, whereas other sentences—those crucially involved in paradoxes—can never achieve absolute generality. The position of bicontextualism will be developed as a position combining the best of two worlds. However, before we can combine the best of these two worlds, we first have to identify it.

To begin, I will emphasize the key virtues of generality absolutism: absolutism is indispensable for successful semantic theorizing, as it provides an essential fallback position that prevents the existence of semantic pariahs—objects that cannot be named in natural language or any possible extension thereof. Furthermore, some of the most fundamental claims in logic, metaphysics, and scientific inquiry express facts of such overarching generality that any less than all-encompassing quantifier domain would result in misinterpretation. Consequently, an absolutist interpretation is required to grant these facts their intended interpretation. These features render absolutism a desirable feature of our semantic theories.

However, generality relativists argue that absolutism is prone to paradoxes, and therefore reject it outright. They argue that the concept of absolute generality is incoherent and that relativism provides a safer alternative. Despite this critique, the argument against absolutism is flawed for two reasons. First, the debate between absolutism and relativism has often been presented as rigid and polarized, leaving no room for intermediate positions. However, the bicontextualist position shows that such a middle ground exists and is a natural outgrowth of the discussion. Second, the argument from paradoxes—that no quantifier can achieve absolute generality—is overly hasty. While the paradoxes imply that the quantifiers of *some* specific sentences must be restricted, they fail to imply that the quantifiers of *all* sentences must be restricted.

Despite the failure of the relativistic argument from paradox, the insights to be drawn from the relativistic solutions are invaluable: in the truth-theoretic setting, contextualist solutions to the Liar paradox highlight the inherently hierarchical nature of the concept 'truth'. In the set-theoretic setting, it is the cumulative-iterative approach to Russell's paradox that shows the inherently hierarchical nature of the concept 'set'. Any reasonable approach to the paradoxes must respect the hierarchical nature of these concepts.

In light to this, I will develop bicontextualism as a hybrid approach that combines the advantages of relativistic and hierarchical solutions to paradoxes

with the strengths of absolutism. This approach offers a balanced middle ground that avoids the extremes of both positions.

I will proceed as follows: First, I will sketch the debate between absolutists and relativists with the aim to extract key features for a viable intermediate position (section 2.1). Next, I will examine the paradoxes and their relativistic and hierarchical solutions in detail. The conclusion of this analysis will be that hierarchical solutions to paradoxes are the most sophisticated and elegant solutions on offer, and that a desirable intermediate position combines absolutism with the elegant hierarchical solutions to paradoxes (section 2.2). After that, I will show how relativists use the paradoxes to argue against absolutism (section 2.3). However, in the final section, I will challenge the relativist conclusion as overly hasty and propose bicontextualism as a natural and compelling intermediate position (section 2.4).

2.1 ABSOLUTISM AND RELATIVISM

I will now introduce *generality absolutism* and *generality relativism*. I will take absolutism to be the default stance due to its intuitive nature, while relativism will be developed as the opposing standpoint. I will not sketch the whole debate between absolutists and relativists, as this has already been done.[1] Instead, I want to highlight what I take to be the main advantages of both positions. The aim of this section is to identify those features of absolutism and relativism that I want preserve in the intermediate position. I begin with absolutism.

The main claim of generality absolutism is this:

> **Generality Absolutism:** It is possible to quantify over absolutely everything. Quantifiers can range unrestrictedly.

According to absolutism, there are some cases in which 'everything' really means everything in the sense of 'absolutely everything there is'. Absolutism can similarly be developed as a semantic thesis in more technical terms. Then, absolutism is the claim that there is a quantifier domain containing absolutely everything, or alternatively, that there is an interpretation I of a given object language $\mathcal{L}$ according to which the $\mathcal{L}$-quantifiers range over absolutely everything.[2] Within this work, I will use different ways of speaking so that 'unrestricted quantification', 'all-inclusive quantifier domain', 'generality absolutism', or simply 'absolutism' all describe the same phenomenon and can be used interchangeably.

Generality absolutism is not the claim that we *always* quantify over absolutely everything. Domain restrictions are unproblematic for the absolutist. Given a quantifier $\forall$ with an all-inclusive domain, absolutists can model domain-restrictions as a restriction to a subdomain. E.g., absolutists can model *D*-

1 See the introduction section of Rayo and Uzquiano [134], Florio [38], and Studd [152, ch. 1].

2 There are actually two different questions here: (1) can we quantify over absolutely everything? And (2) Is there a quantifier domain containing absolutely everything? However, taking a liberal views regarding what counts as a domain as I will do, (1) and (2) coincide.

quantification, where D is not universal, via the *universal* universal quantifier $\forall$ and a predicate D:

$$\forall x(D(x) \rightarrow ...).$$

Hence, absolutism can be interpreted as the claim that it is possible to quantify over absolutely everything, *sometimes*. Bicontextualism, ultimately, can also be summarized as the claim that unrestricted quantification is *sometimes* available, even though bicontextualists interpret *sometimes* quite differently from the absolutist (I will come back to that at the end of this section).

The main claim of generality relativism is this:

> **Generality Relativism:** It is impossible to quantify over absolutely everything. Quantifiers are always restricted.

According to generality relativism, there are no cases in which 'everything' really means 'absolutely everything there is', and quantifier ranges are always restricted to a less than all-inclusive subcollection of absolutely everything.[3] Quantifiers can be restricted in many different ways, e.g., by explicit syntactic restrictions as in 'every donkey', or by implicit restrictions due to contextual limitations.[4] As in the case of absolutism, I will use different ways of speaking about relativism, so that 'less than all-inclusive quantifier domain', 'restricted quantification', 'generality relativism', or simply 'relativism' all describe the same phenomenon and can be used interchangeably.

The position of generality relativism is notoriously difficult to state. The above description of generality relativism is self-defeating since it involves the quantifier 'absolutely everything', and only when we assume that this quantifier ranges over absolutely everything, we succeed in giving a proper description of relativism. The literature concerning this phenomenon is vast. In my view, the most promising approaches to express generality relativism are *schematic generality* and *modal generality*. Proponents of the former approach formulate relativism not as a single maximally general statement, but as an open-ended schema, whose infinitely many instances deny the comprehensiveness of particular domains or universes, with each instance framed from the perspective of a potentially more inclusive domain or interpretation. The later uses suitably interpreted modal operator and treats quantifiers as *modalized quantifiers*, ranging over domains that can expand across possible contexts or stages. Key to this approach is to take the modal operators as primitive, because giving a

3 Strictly speaking, this is only true for the *restrictionist* version of relativism, but not for the *expansionsist* version (see Studd [152, ch. 4]). The difference is the following: let U be the *universe of discourse* which contains absolutely everything, and let D be the domain of the particular quantifiers that we're considering. A restrictionist claims that in every context, quantifier domains are less than all-inclusive subcollections of the universe, i.e. $D \subset U$, and each domain D is open to expansion but will never exhaust the universe U. Expansionists, on the other hand, claim that it is possible to have quantifiers range over the entire universe U, but every universe is open for expansion. However, if we use a strict reading of absolutism as the claim that we can quantify over everything *there is*, then expansionism is, in my view, closer to absolutism than to relativism. I will therefore focus on the restrictionist flavor of relativism.

4 As we will see shortly, Williamson [169] has argued that even such seemingly innocent cases as 'every donkey' presuppose unrestricted quantification.

semantics for modalized quantifier expressions such as a possible world semantics would involve quantification over all models.[5] Whether these attempts to express generality relativism are successful is controversial. However, as McGee points out, "the fact that a thesis cannot be coherently maintained does not strictly entail that the thesis is false" [100, p. 55], and I believe that, despite these difficulties, the position is clear enough so that we can work with it.

The strongest argument for relativism is an argument from paradoxes. In the simplest form, the argument runs as follows: domains of discourse are sets, and so an absolutist requires, as the domain of their unrestricted quantifier, a universal set. However, there cannot be a universal set on pain of contradiction, and hence, there cannot be a quantifier domain in the sense of absolutism. The crucial step involves Russell's paradox to show that for any allegedly universal set U, we can construct U's Russell set, R_U, as the set of all non-self-membered sets of U. This leads to a paradox if we assume that R_U is in U, and hence, we conclude that R_U is not in U. But if this is true, U is not universal. I will discuss this argument at length in section 2.2, so that I can set is aside for now.

Let me now outline what I believe are the key advantages of absolutism. In a nutshell, absolutism offers a much more coherent picture of (natural) language with greater expressive power then relativism; this concerns so-called *kind generalization* and the scope of prima facie absolutistic claims. Moreover, semantic theorizing crucially depends on unrestricted quantification because only absolutism allows for unambiguous definitions of key-concepts such as logical consequence and the characterization of logical validities. On the other hand, relativism has some difficulties regarding the open-ended character of natural language and the potential ambiguity of logical validities.

Let's begin with kind generalizations: Consider the sentences 'every donkey is mulish'. The ability to make such generalizations is one of the most fundamental features of language, yet, as Williamson [169] argues, even these seemingly restricted generalizations presuppose an absolutist notion of quantification. The sentence 'Every donkey is mulish' is true if and only if *absolutely every donkey* is mulish. How can the relativist express this, given that their resources comprise only restricted domains? For this sentence to be true under the relativist interpretation, there must be a relativistically acceptable domain M that includes all donkeys. But how can the relativist specify such a domain? A straightforward specification of M is $\forall x(Dx \to \exists_M y(y = x))$. This guarantees that every donkey is indeed in M. But this is a guarantee that purely relativistic resources cannot provide on their own. Without a relativistically acceptable specification of M, it remains unclear from the relativists' point of view whether their generalizations genuinely capture the scope of such statements. Absolutists face no such challenges, as they can rely on their all-inclusive domain as a fallback position. If this is true, relativists either have to employ resources

5 Note that both, schematic and modal generality, provide ways to simulate a form of generality that is not quantificational. While schematic generality is limited to Π_1-sentences, modal generality can provide non-quantificational generalizations for sentences of arbitrary complexity. I will compare bicontextualism with schematic and modal generality in section 6.3.2. See Lavine [81] and Studd [152] as well as some applications of these approaches to set theory and the notion of infinity as in Studd [154], Linnebo [90], Linnebo and Shapiro [94], and Berry [10]. Some of these ideas can already be found in Parsons [116].

tantamount to absolutism to articulate the needed domain restrictions, or they cannot properly express kind generalizations.[6]

Let's now consider prima facie absolutistic claims: Many claims require an unrestricted quantifier domain in order to express their intended meaning. These are claims which, when interpreted over a less than all-inclusive domain, either fail completely to express what they intend to express and turn into trivialities, or even come dangerously close to falsehoods. The most obvious examples come from logic and metaphysics. Consider the claim

1. Everything is self-identical ($\forall x\ x = x$).

According to the intended interpretation, 1 says that absolutely everything is self-identical. In order to capture the intended interpretation, 1 requires an unrestricted quantifier domain. This is due to the fact that the reflexivity of identity is considered one of the most fundamental and universal principles of logic, and that it applies to everything whatsoever.[7]

Further examples are found in the foundations of mathematics:

2. Nothing is an element of the empty set ($\neg\exists x\ x \in \emptyset$).
3. Two sets are identical if and only if they have the same elements ($\forall x \forall y(\forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y)$).

2 requires an all-inclusive domain for its intended meaning. Even though set theory mostly operates without *urelements*, having a domain containing just all sets is not sufficient here (and even that would require absolutism according to (1)). The empty set cannot contain anything, neither sets nor urelements. This is for conceptual reasons. Moreover, it's not clear how to interpret 2 relativistically. Taking relativism at face value, 2 only says that with respect to the current domain D, there is nothing in D that is contained in the empty set. What relativists cannot say, however, is that no domain whatsoever contains something that is an element of the empty set, because quantification over all domains is tantamount to relativism, given that everything is contained in some domain.

The axiom of extensionality (3), giving an identity criterion for sets, is perhaps the most fundamental principle in modern set theory. Many have argued that extensionality is so basic for our modern conception of sets that if anything does not satisfy it, it is simply not a set.[8] So, just as 2, the axiom of extensionality requires a domain in the sense of absolutism for its intended interpretation.

The debate about absolutism and relativism has far-reaching consequences for the definition of key concepts in formal semantics, such as the notion of

6 Note that expansionists have a way to respond to the objection from kind generalizations, because they can rely on the background universe and express 'Every donkey is mulish', just like the absolutist, as $\forall x(Dx \rightarrow \exists_M y(y = x))$, with $\forall$ ranging over the whole universe. However, in my view, this is only because expansionists are closer to absolutism than to relativism. See Studd [152, ch. 4] and also footnote 3 of this chapter.

7 Similar considerations apply to any universal property. Consider, for example, the central claim of physicalism, which is that "Everything is physical" Stoljar [151]. Universal claims remain true when interpreted over a less than all-inclusive domain—if everything is physical, any proper subcollection contains only physical things. But when interpreted over a restricted domain, this claim fails to express its intended meaning.

8 See Boolos [15] and Maddy [97], or more recently Incurvati [73, ch. 1].

logical consequence. Consider the standard definition of logical consequence from Ebbinghaus, Flum, and Thomas's *Mathematical logic*:

> φ is a consequence of Φ (written $\Phi \models \varphi$) iff *every interpretation* which is a model of Φ is also a model of φ. [32, p. 31, my italics].

The question immediately arises how to interpret quantification over every interpretation in the metalanguage. Only if the metalinguistic quantifier truly ranges over every interpretation, the definition of logical consequence captures it's intended meaning as truth preservation over *every interpretation*. But under the plausible assumption that everything is contained in some domain, the metalinguistic quantifier in the definition of logical consequence is de facto absolutistic—that is, quantification over every interpretation implies quantification over absolutely everything. Conversely, if there are objects not contained in any interpretation—so-called *semantic pariahs*—the quantifier is de facto *not* absolutistic.

Semantic pariahs, by their very nature, are objects about which we cannot speak in any language. Colloquial English has the ability to name every object whatsoever, at least in principal. Even if the language we currently speak cannot name everything, it is at least possible, for every entity, to expand the language to include a name or term that refers to this entity. This *open-endedness* is a core feature of natural language.[9] So it seems that semantic pariahs do not exist. But if they do not exist, the metalinguistic quantifier, according to the relativist, cannot range over every interpretation as this would make it absolutistic. But if the metalinguistic quantifier is restricted, the notion of logical consequence is only defined as truth preservation over *some interpretations*, i.e., as truth preservation over the models within the domain of the metalinguistic quantifiers. This makes the definition of semantic key-concepts potentially ambiguous. In short, relativism implies the following disjunction: either the metalinguistic quantifier has restricted scope, but then the definition of logical consequence is potentially ambiguous, or the metalinguistic quantifier has unrestricted scope, but then semantic pariahs exist. Absolutism faces none of these problems.

The ambiguity of the definition of logical leads to a potential ambiguity of the notion of logical validity. When quantifier domains are always restricted, logical validities can never be asserted unconditionally. This stands in stark contrast to the intuitive notion that certain logical truths, such as $(A \to A)$ or $\forall x(P(x) \vee \neg P(x))$, express fundamental facts that should hold universally, without any restrictions whatsoever. However, according to the relativist, $\forall x(P(x) \vee \neg P(x))$ can only mean 'Everything *in the current domain* is either P or not-P', and so the meaning of such statements changes from context to context. As Glanzberg puts it, all relativists can achieve is "persistence: insensitivity to expansion of the domain" [49, p. 560]. Whether and how this form of persistence can be expressed in a relativistic framework remains controversial.[10]

9 See McGee [100, 101] and Williamson [169].
10 See Glanzberg [49] and Parsons [119].

Let me summarize: absolutism is the claim that unrestricted quantification is possible, relativism is that claim that it is not possible. The positive feature of absolutism are mostly connected with its crucial role for semantic theorizing. Without absolutism, even mild forms of generalization (kind generalization) are at risk, and prima facie absolutistic claims loose much of their force. Moreover, relativism has some disadvantages concerning semantic pariahs and ambiguities of key-concepts of semantics. As soon as we have absolutism as a fall-back position, the disadvantages of relativism vanish, and the advantages of absolutism come into force.

The intermediate position that I'm about to develop shall have all these features, i.e. it shall be absolutistic enough to allow kind generalizations, give prima facie absolutistic claims their intended interpretation, and avoid the disadvantages of relativism. The rest of this section deals with the strongest argument for relativism—the avoidance of set-theoretic and semantic paradoxes—and how to combine the elegant relativistic approaches with absolutism.

2.2 SOME PARADOXES AND THEIR (HIERARCHICAL) SOLUTIONS

The strongest argument for relativism is that absolutism is prone to paradoxes. In this section, I will analyze two of the most fundamental paradoxes, *Russell's paradox* and the *Liar paradox*, along with a generalization of Russell's paradox due to Timothy Williamson. I will argue that both semantic and set-theoretic paradoxes are, at their core, instances of *diagonal arguments*—arguments that allow one to *diagonalize out* of an initially given totality and identify certain objects such as sets, truths, or interpretations, depending on the context that that are not part of the initial totality, thereby revealing the existence of levels of a hierarchy of some kind. Relativists use this fact to challenge generality absolutism: the hierarchical nature of the concepts of 'set' and 'truth' precludes, so they argue, the possibility of absolute generality for quantifiers.

A characteristic feature of paradoxes is that they begin with seemingly unproblematic assumptions and proceed by seemingly unproblematic inferences to an evidently false conclusion. The paradigmatic instances of such assumptions concern the naïve notions of 'set' and 'truth'. Relativists have developed sophisticated responses to these paradoxes, offering deep philosophical insights by developing alternative, non-naïve, and explicitly *hierarchical* accounts of these concepts.

This brings me to a further aim of this section. I do not wish to neglect these insights. Rather, the intermediate position I will develop seeks to preserve the hierarchical solutions to the paradoxes.

2.2.1 *Russell's Paradox*

In this section, I analyze Russell's paradox and its most prominent solution: the *iterative conception of set* (together with its ontological counterpart, the *cumulative hierarchy*). Russell's paradox is certainly one of the most famous paradoxes in the history of (modern) logic, and its impact on the development of modern mathematics cannot be overestimated. The paradox is a result of the *naïve*

conception of set according to which each property φ determines a set containing all elements that satisfy φ. So, for example, if the property is 'is a natural number', then by the naïve conception there exists a set of all natural numbers. The naïve conception was taken to be a self-evident principle of set existence, and the fact that it is now called *naïve* is due to the paradoxes that have proved it wrong. Informally, Russell's paradox has the following form. Start from the naïve conception and consider the set R of all sets that are not members of themselves. Now ask: is R itself a member of R? If R is a member of R, then by the definition of R it must be non-self-membered, and hence R is not a member of R. On the other hand, if R is not a member of R, then it is not self-membered, so it satisfies the condition for being in R, and hence R is a member of itself. This means that R is a member of itself if and only if R is not a member of itself. I will now make this informal presentation more precise and rephrase it in the terminology of modern set theory.

Important for the derivation of Russell's paradox are the following ingredients: the principle of naïve comprehension and classical logic. The principle of naïve comprehension, understood as a set-theoretic principle, is the formal counterpart of the naïve conception of set. It states that for every $\mathcal{L}_\in$-definable property $\varphi(x)$, there exists a set y containing all and only the elements that satisfy $\varphi(x)$, i.e. every predicate has a set as its extension. Formally, naïve comprehension is an axiom schema, consisting of infinitely many axioms:

> **Naïve Comprehension:** For every $\mathcal{L}_\in$-formula $\varphi(x)$, the following is an axiom:
>
> $$\exists y \forall x (x \in y \leftrightarrow \varphi(x)).$$

Note that the existence of a universal set, which plays an important role for generality absolutism, can be deduced from naïve comprehension as follows: let $\varphi(x) \equiv x = x$. By naïve comprehension, there is a set $y = \{x \mid x = x\}$ which is universal by the reflexivity of identity.

The most straightforward way to derive Russell's paradox follows the above template and starts with $\varphi(x) \equiv x \notin x$. Then, by naïve comprehension, there exists a set $R = \{x \mid x \notin x\}$. Now, if $R \in R$, it must satisfy the defining property of R which is $R \notin R$. If, on the other hand, $R \notin R$ it satisfies the defining property of R, and so $R \in R$. So from naïve comprehension alone, it follows that $R \in R \leftrightarrow R \notin R$.

At this point, we have two different assumptions: naïve comprehension and classical logic. Some have tried to refute classical logic and to uphold the naïve conception. Most famously, Graham Priest [126] develops naïve set theory based on paraconsistent logic.[11] Another approach to rescue the naïve conception is due to Quine, who believed that the naïve conception is the only intuitive conception of set. Quine's theory *New Foundations* [127] introduces a syntactic criterion (of stratification) to single out *good* instances of naïve comprehension from *bad* instances.[12] The vast majority of philosophers and mathematicians,

11 See also Badia, Weber, and Girard [4], Restall [140], Weir [167], and Weber [165, 164], as well as Incurvati [73, Ch. 4] for an overview.

12 See Morris [105] and Incurvati [73, ch. 6] for details.

however, refuted the naïve conception, and replaced it with the nowadays standard *iterative conception of set*.[13] In comparison, the iterative conception can best be described as a *bottom-up* approach to the set-concept, while the naïve conception is *flat*: the naïve, flat approach assumes that the totality of all things (sets) is given at the outset—think of all these things as spread out in a flat, two dimensional plane—and definable properties divide this totality in an extension and an antiextension of the predicate. The iterative, bottom-up approach, on the other hand, starts with no previously given sets (i.e. only with non-sets that are called *urelements*, or with no objects at all, i.e., only with the empty set $\emptyset$), and constructs new sets stepwise out of these objects by iterated application of the set-construction methods that are encoded in the axioms of modern set theory. Gödel sums this up nicely in the following famous passage of his paper on the *continuum hypothesis*:

> This concept of set, however, according to which a set is something obtainable from the integers (or some other well-defined objects) by iterated application of the operation "set of", not something obtained by dividing the totality of all existing things into two categories, has never led to any antinomy whatsoever; that is, the perfectly "naïve" and uncritical working with this concept of set has so far proved completely self-consistent. [53, pp. 474–5]

The nowadays standard system of *Zermelo-Fraenkel Set Theory with the Axiom of Choice* (ZFC) is an axiomatization of the iterative conception.[14] It keeps the extensionality principle from the naïve conception, according to which two sets are identical if and only if they have the same elements, but replaces the naïve comprehension axioms with the more restricted separation principle, together with further set existence axioms. Formally, separation is an axiom schema, consisting of infinitely many axioms:

> **Separation:** For every $\mathcal{L}_\in$-formula $\varphi(x)$, the following is an axiom:
>
> $$\forall z \exists y \forall x (x \in y \leftrightarrow x \in z \wedge \varphi(x)).$$

Unlike naïve comprehension, separation sanctions set-existence only for subsets of previously given sets.[15]

In ZFC, the derivation of Russell's paradox is blocked: Let M be a set and consider its Russell set R_M via the formula $\varphi \equiv x \notin x$. This yields the following instance of separation:

$$\exists y \forall x (x \in y \leftrightarrow x \in M \wedge x \notin x)$$

which defines R_M as $\{x \in M \mid x \notin x\}$. If we now ask whether $R_M \in R_M$, we arrive at a paradox that $R_M \in R_M \leftrightarrow R_M \notin R_M$ by exactly the same reasoning.

13 See, e.g., Boolos [15], Incurvati [73], and Potter [125].

14 A full presentation of the axioms of set theory is given in section 3.3.

15 Unlike naïve comprehension, separation is a relative set existence axiom. This means that separation alone is not enough to prove the existence of any set. In order to get started, ZFC adds certain absolute set existence axioms which postulate the sets that are not subject to any preconditions such as the *Axiom of Infinity*. More on that in section 3.3.

This time, however, there is a way out. The existence of R_M is subject to two conditions: $x \notin x$ *and* $x \in M$. So, in order to ask whether $R_M \in R_M$, we have implicitly assumed that $R_M \in M$. The strategy to avoid the paradox is to reject the second assumption: $R_M \notin M$—this blocks the paradox outright because if $R_M \notin M$, then R_M cannot be in R_M (since R_M is by definition a subset of M).

By the reasoning of Russell's paradox, we have identified, for an arbitrary set M, a set R_M, which lies outside of M. If we assume that there is a universal set, we arrive again at an inconsistency: let U be the universal set and consider $R_U = \{x \in U \mid x \notin x\}$. We again reason along the line of Russell's paradox and conclude, to avoid the paradox, that $R_U \notin U$, and therefore, there exists a set outside of U. However, since U is universal, this move is not available for conceptual reasons, because for the universal set U, there simply cannot be sets outside of U. In other words, separation blocks the paradox, but the additional assumption of a universal set leads again to paradox. Nowadays, Russell's paradox is interpreted as showing that ZFC, unless inconsistent, cannot prove the existence of a universal set. This fact is so deeply rooted in our current understanding of the set concept that we find *reductio ad absurdum* proofs of the non-existence of a universal set right at the beginning of any standard textbooks.[16]

The development of ZFC was a milestone for modern set theory, and appeared first, in a suitably adjusted second-order version, in Zermelo [173].[17] In this paper, Ernst Zermelo not only gave the first full formulation of ZFC, he also made a strong case for the view of the set-theoretic universe (V) as a *cumulative hierarchy of sets* with its characteristic V-shape.[18] The main idea of the cumulative hierarchy is that sets are arranged in stages which are indexed by ordinals. At stage 0, the cumulative hierarchy starts with the empty set.[19] Successor stages are given by the powerset of the previous stage, and limit stages are given by taking the union of all previous stages. Formally, the cumulative hierarchy is defined by *transfinite recursion* as follows:

$$V_0 := \emptyset, \quad V_{\alpha+1} := \mathcal{P}(V_\alpha), \quad V_\gamma := \bigcup_{\alpha<\gamma} V_\alpha \quad \text{(for limit ordinals } \gamma\text{)}.$$

16 See Kunen [80, p. 21, lemma I.4.9] or Jech [75, p. 8].

17 The second-order version ZFC_2 replaces the first-order schemes of Replacement and Separation by second-order axioms. The main difference between ZFC and ZFC_2 is that ZFC_2 is a *quasi-categorical* axiomatization of the iterative conception, yielding only *structurally nice* models. More on that in section 3.3.

18 The idea of the cumulative hierarchy can already be found in an earlier publication by John von Neumann [111]. However, von Neumann's axiomatization does not include the Axioms of Replacement and Choice. Moreover, unlike nowadays standard, von Neumann took the notion of a function, and not the notion of a set, as the starting point (Ausgangspunkt) of his work. See Kanamori [76, p. 518] for details.

19 Strictly speaking, in Zermelo's presentation we start at stage 0 with a collection of urelements which are non-sets (i.e. urelements do not have members), and which build the starting block of the construction. In the Gödel-quote given above, these are the "integers (or some other well-defined objects)". For purely mathematical purposes, however, urelements are dispensable as mathematical objects can always be mimicked by sets. For instance, we can mimic the natural numbers $\mathbb{N}$ by *von-Neumann ordinals* and then construct all other number systems such as $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ from the set-theoretic surrogate of $\mathbb{N}$. Because of this, I will ignore urelements for the rest of the presentation.

The cumulative hierarchy, V, is defined as the union of all these ranks:

$$\mathsf{V} := \bigcup_{\alpha \in \mathrm{Ord}} V_\alpha.$$

Zermelo was able to show that models of (the second-order version of) ZFC are exactly those layers of V that satisfy certain closure conditions. These layers are called *inaccessible ranks of the cumulative hierarchy*.[20] As Kanamori observes, the impact of Zermelo's work on models of set theory was so strong that the cumulative hierarchy is nowadays considered "as a striking, synthetic view of a procession of natural models" [76, p. 521]. By it's construction, the cumulative hierarchy perfectly encapsulates the iterative conception of set. As described above, we start at stage 0 with no previously given objects, and so the starting set is empty. We then iterate the application of set-of operations in form of the powerset, and the union operation. The resulting picture is that we have a philosophical conception of set (the iterative conception), axiomatized by a strong theory (ZFC), and a picture of the set-theoretic universe as a cumulative hierarchy which contains the *standard model* of the axiomatization (ZFC) of the iterative conception.

The treatment of Russell's paradox is straightforward: consider any inaccessible rank V_κ in the cumulative hierarchy (i.e. any standard model of ZFC), and take any set set $a \in V_\kappa$. Then, a appears for the first time at some layer V_α below V_κ. For simplicity, let us assume that a is identical to this layer V_α. By the axiom of separation, we can now consider the Russell set $R_\alpha = \{x \in V_\alpha \mid x \notin x\}$ of V_α. Since R_α is not in V_α, R_α can not be in R_α. This blocks the contradiction outright. Since R_α is not an element of V_α, it must enter the cumulative hierarchy only at a later stage. And indeed, since R_α contains only elements of V_α but is not itself an element of V_α, it must be a subset of V_α. R_α is therefore an element of $\mathcal{P}(V_\alpha)$ (the collection of all subsets of V_α) which is $V_{\alpha+1}$. Considering the Russell set of V_α therefore leads us naturally to the subsequent stage $V_{\alpha+1}$. This reasoning can be applied at any stage V_β (for $\beta < \kappa$ and ordinal) inside V_κ: by the above mentioned closure properties, V_κ contains all the V_β's, all the $\mathcal{P}(V_\beta)$'s, and all the union at limit stages γ (for $\beta < \gamma < \kappa$).

We can now ask what about V_κ and its Russell set $R_\kappa = \{x \in V_\kappa \mid x \notin x\}$? This is where Zermelo's picture of the open-endedness of the cumulative hierarchy unleashes its full power: V_κ does not contain V_κ as a set, and so we cannot ask *inside* V_κ for R_κ. However, since V is open-ended, there is a next inaccessible rank V_μ, that is also a model of ZFC and satisfies the same closure properties as V_κ. We can therefore consider any Russell set for any set $b \in V_\mu$. Since V_κ is one of these b's, we therefore find R_κ in V_μ precisely at stage $V_{\kappa+1}$. More generally, for every stage V_α in V, the Russell set R_α of V_α enters the cumulative hierarchy at the successor stage of V_α. Since V is open-ended, every stage has a successor stage (and every infinite sequence of successor stages is followed by a limit stage which is the union of all previous stages). This, however, shows the hierarchical nature of the modern solution of the paradoxes. We can think

20 The requirement is closure under set construction corresponding to the available axioms. For example, since the *Axiom of Powerset* is one of these axiom, a model of the theory has to contain, for every set A, its powerset $\mathcal{P}(A)$, and so on for all other axioms. The inaccessible ranks are exactly defined to meet these conditions. See section 3.3 for details.

of the cumulative hierarchy as a sequence of models where Russell sets come up as follows:

$$\begin{array}{ccccccccccccc} \ldots & V_\alpha & \in & V_{\alpha+1} & \in & V_{\alpha+2} & \in & \ldots & \in & V_\gamma & \in & V_{\gamma+1} & \ldots \\ & \cup\!| & \not\in & \cup\!| & \not\in & \cup\!| & \not\in & & & \cup\!| & \not\in & \cup\!| & \not\in \\ \ldots & R_\alpha & & R_{\alpha+1} & & R_{\alpha+2} & & \ldots & & R_\gamma & & R_{\gamma+1} & \ldots \end{array}$$

In short, the cumulative hierarchy solution to Russell's paradox is hierarchical. The Russell sets *diagonalizes out* of any given set and enters the hierarchy at the next layer. By the open-endedness of the cumulative hierarchy, this process can be iterated indefinitely, so that we can always find further layers of V.

2.2.2 *The Liar Paradox*

In this section, I analyze the Liar paradox and one of its most prominent solutions: *contextualism*.[21] I focus here on *type-free* (i.e. self-applicable) truth à la Kripke [79] and *contextualism* à la Parsons [118] and Glanzberg [50, 47, 51] because they combine many desirable features in a natural way: a type-free and self-applicable notion of 'truth' that can model natural language phenomena such as iterations of the truth predicate ('"'φ' is true' is true') and blind ascriptions ('Everything Kripke says about truth is true').

Contextualism, more generally, incorporates the idea that the meaning of linguistic expressions depends on the context of utterance. Regarding the Liar sentence, the Parsons-Glanzberg line of contextualism holds that there is no proposition corresponding to the Liar sentence in the context of utterance, and only later contexts contain such propositions.[22]

The Liar paradox can be presented in different forms, for instance as 'This sentence is false', 'This sentence is not true', etc. Taking 'This sentence is not true' to be the standard version of the Liar sentence, the paradox results from asking whether the Liar sentence is true or not. If the Liar sentence is true, then it is true that the Liar sentence is not true, and hence, the Liar sentence is not true. If the Liar sentence is not true, then, it is not true that the Liar sentence is not true, and hence, the Liar sentence is true. This means that the Liar sentence is true if and only if the Liar sentence is not true. I will now make this informal version of the argument more precise.

Important for the derivation of the paradox are the following three ingredients: a truth predicate that validates the naïve truth principles, classical logic, and some form of self-reference.

21 For general background on contextualist approaches to the semantic paradoxes, see Parsons [118], Glanzberg [50, 47], and Murzi and Rossi [109, 110].

22 The central result of this section will be that contextualist solutions to the Liar paradox are hierarchical. The most prominent alternative solution to the Liar paradox from the seminal paper Tarski [156] (1933 in polish) is similarly hierarchical. Unlike the solutions presented here, Tarski suggests to construct the truth predicate, for a given object language $\mathcal{L}$, in a metalanguage $\mathcal{L}'$, so that truth-in-$\mathcal{L}$ is a $\mathcal{L}'$-predicate which applies only to $\mathcal{L}$-sentences. In order to model truth-in-$\mathcal{L}'$, however, we have to go to yet another metalanguage $\mathcal{L}''$ to construct truth-in-$\mathcal{L}'$ as a $\mathcal{L}''$-predicate which applies only to $\mathcal{L}'$-sentences. This process can be iterated to an open-ended successions of languages and metalanguages. As all the metalanguages are strict extensions of the object language, this solution is obviously hierarchical. For more, see Halbach [57, chs. 3, 7, 8, 9].

As usual in the literature on theories of truth, I work with an arithmetical base language that I use as a toy model of natural language.[23] I add an axiomatic theory building on *Dedekind-Peano Arithmetic* (PA) which can define a sufficiently strong coding schema, $\lfloor \cdot \rfloor$, that can express its own syntax, and that can prove the strong diagonalization lemma that is required for self-referential sentences.[24] The resulting schema will be such that, for any $\mathcal{L}_{\mathsf{Tr}}$-sentence φ, $\lfloor \varphi \rfloor$ is the name of φ. In this system, we can define the Liar sentence λ as "'λ' is not true', formally $\lambda :\leftrightarrow \neg \mathsf{Tr}(\lfloor \lambda \rfloor)$.

To express the naïve truth principles, I expand the language further by adding a fresh, unary predicate constant Tr for truth. Call the resulting language $\mathcal{L}_{\mathsf{Tr}}$. The naïve truth principles state that φ and $\mathsf{Tr}(\lfloor \varphi \rfloor)$ are always interderivable, i.e. we assume $\varphi \leftrightarrow \mathsf{Tr}(\lfloor \varphi \rfloor)$ for every $\mathcal{L}_{\mathsf{Tr}}$-sentence φ. The idea is that whenever we can infer a sentence φ, we can reason from φ to "'φ' is true', i.e., $\mathsf{Tr}(\lfloor \varphi \rfloor)$ and *vice versa*. Some call these rules *capture* (φ implies $\mathsf{Tr}(\lfloor \varphi \rfloor)$) and *release* ($\mathsf{Tr}(\lfloor \varphi \rfloor)$ implies φ).[25] Other formulate these principles as inference rules and call them introduction (Tr-I) and elimination (Tr-E) rules:

$$\mathsf{Tr\text{-}I}\ \frac{\Gamma \vdash \varphi}{\Gamma \vdash \mathsf{Tr}(\lfloor \varphi \rfloor)} \qquad\qquad \frac{\Gamma \vdash \mathsf{Tr}(\lfloor \varphi \rfloor)}{\Gamma \vdash \varphi}\ \mathsf{Tr\text{-}E}$$

Either way, the important principle is that the base system (PA) together with the principles governing the truth predicate also contains all Tarski-biconditionals of the form $\varphi \leftrightarrow \mathsf{Tr}(\lfloor \varphi \rfloor)$. It is now routine to derive the paradox and even to show that the system is trivial:[26]

1. $\mathsf{Tr}(\lfloor \lambda \rfloor) \vdash \mathsf{Tr}(\lfloor \lambda \rfloor) \wedge \neg \mathsf{Tr}(\lfloor \lambda \rfloor)$:
   - (a) $\mathsf{Tr}(\lfloor \lambda \rfloor)$ (assumption)
   - (b) λ (Tr-E)
   - (c) $\neg \mathsf{Tr}(\lfloor \lambda \rfloor)$ (def. λ)
   - (d) $\mathsf{Tr}(\lfloor \lambda \rfloor) \wedge \neg \mathsf{Tr}(\lfloor \lambda \rfloor)$ ((a),(c), $\wedge$-I)
2. $\neg \mathsf{Tr}(\lfloor \lambda \rfloor) \vdash \mathsf{Tr}(\lfloor \lambda \rfloor) \wedge \neg \mathsf{Tr}(\lfloor \lambda \rfloor)$:
   - (a) $\neg \mathsf{Tr}(\lfloor \lambda \rfloor)$ (assumption)
   - (b) λ (def. λ)
   - (c) $\mathsf{Tr}(\lfloor \lambda \rfloor)$ (Tr-I)
   - (d) $\mathsf{Tr}(\lfloor \lambda \rfloor) \wedge \neg \mathsf{Tr}(\lfloor \lambda \rfloor)$ ((a), (c), $\wedge$-I)
3. $\mathsf{Tr}(\lfloor \lambda \rfloor) \vee \neg \mathsf{Tr}(\lfloor \lambda \rfloor) \vdash \mathsf{Tr}(\lfloor \lambda \rfloor) \wedge \neg \mathsf{Tr}(\lfloor \lambda \rfloor)$ (1., 2.)
4. $\mathsf{Tr}(\lfloor \lambda \rfloor) \wedge \neg \mathsf{Tr}(\lfloor \lambda \rfloor)$ (3., LEM)
5. ψ (4., EFQ)

23 Some authors recently work with a syntax theory instead of an arithmetical base theory (see Mount and Waxman [108], Fujimoto [44], Leigh and Nicolai [82], and Nicolai [113, 112]). However, the latter is still the predominant approach. See for instance Halbach [57], Beall, Glanzberg, and Ripley [7], Field [34], McGee [102], Cieśliński [22], and Horsten [71].

24 The formal details are given in section 3.2.1. I use $\lfloor \cdot \rfloor$ instead of the more convenient $\ulcorner \cdot \urcorner$ because this is already in use (see section 3.4).

25 See Beall, Glanzberg, and Ripley [7].

26 My presentation follows Beall, Glanzberg, and Ripley [7, ch. 3].

The principle of classical logic that we have used are conjunction introduction ($\wedge$-I), the law of excluded middle (LEM), and explosion or ex falso quodlibet (EFQ). Paracomplete and paraconsistent logicians have developed approaches to the Liar paradox that reject LEM and EFQ respectively. However, the costs of these solutions are high as they give up classical logic.

Naturally, the next principle to reject in order to avoid the contradiction is the *naïveté* of truth. One such approach is contextualism.[27] As a first step, contextualists reformulate the Liar paradox by introducing two things: a context parameter, and an ontology of propositions.[28] As a consequence, the Liar sentence λ is no longer the sentence "'λ' is not true', but rather 'There is no proposition p such that 'λ' expresses p in context c and p is true in c'. Let me now make this approach more precise.

Let $\mathsf{Exp}(x, y, z)$ be the three-place predicate 'x expresses proposition y in context z', and *contextualize* the Liar sentence and the truth predicate. $\mathsf{Tr}(\lfloor\varphi\rfloor)$ is now $\mathsf{Exp}(\lfloor\varphi\rfloor, p, c) \wedge \mathsf{Tr}_c(\lfloor p\rfloor)$, i.e. "'$\varphi$' is true' now means that φ expresses a proposition p in context c, and p is true in c. *Naïveté* of truth has to be restricted to sentences that express propositions, i.e., instead of $\varphi \leftrightarrow \mathsf{Tr}(\lfloor\varphi\rfloor)$ we have

$$\mathsf{Exp}(\lfloor\varphi\rfloor, p, c) \rightarrow (\mathsf{Tr}_c(\lfloor p\rfloor) \leftrightarrow \varphi). \tag{2.1}$$

If φ expresses a true proposition p in context c, than we can always infer $\mathsf{Tr}_c(\lfloor p\rfloor)$ from φ and *vice versa*. Contextualists assume that whenever a sentence φ is provable from true premises, then φ expresses a proposition:

$$\vdash \varphi \rightarrow \exists_c p(\mathsf{Exp}(\lfloor\varphi\rfloor, p, c)). \tag{2.2}$$

Moreover, contextualists assume that this proposition is unique:[29]

$$\mathsf{Exp}(\lfloor\varphi\rfloor, p, c) \wedge \mathsf{Exp}(\lfloor\varphi\rfloor, q, c) \rightarrow p = q. \tag{2.3}$$

Contextualists define the Liar of context c as the statement that there is no true proposition expressed by λ in context c. It will be important later on to keep track of the context of utterance of the Liar sentence. This means that every context has it's own Liar sentence:

$$\lambda_c :\leftrightarrow \neg\exists_c p(\mathsf{Ext}(\lfloor\lambda\rfloor, p, c) \wedge \mathsf{Tr}_c(p)). \tag{2.4}$$

The paradoxical Liar argument now runs as follows: First, we show that the assumption that λ_c expresses a proposition in c leads to a contradiction, i.e.

27 The presentation of contextualism here and throughout this work mainly follows Glanzberg [50, 47] and Parsons [118], but uses some of the simplifications developed by Rossi [144]. The derivation of the contextualized liar paradox follows Glanzberg [50].

28 As pointed out already in Parsons [118], it is not necessary to rely on an ontology of propositions. Such entities are notoriously difficult to understand and the word 'proposition' might even lack uniform meaning [72]. I will therefore indicate how to reformulate the contextualist reasoning without these dubious entities. For the narrative, however, I find propositions a helpful heuristics and will therefore continue to use them in the main text.

29 This principle is certainly debatable in a non-contextualist setting. E.g., the sentence 'I am hungry' can certainly mean a lot of different things when uttered by different speakers in different situations. However, in a contextualist setting, this sentence becomes something like 'Person X is hungry in context c', where the uniqueness claims seems far less troubling.

$\exists_c p\mathsf{Ext}(\lfloor\lambda_c\rfloor, p, c) \vdash \mathsf{Tr}_c(p) \wedge \neg\mathsf{Tr}_c(p)$ (steps 1. – (b.iv)). In the second step, we refute that assumption and show that this also leads to a contradiction, now without any assumptions, i.e. $\vdash \exists_c p\mathsf{Ext}(\lfloor\lambda_c\rfloor, p, c) \wedge \neg\exists_c p\mathsf{Ext}(\lfloor\lambda_c\rfloor, p, c)$ (steps 2. – 6.):

1.	$\exists_c p\mathsf{Exp}(\lfloor\lambda_c\rfloor, p, c)$			(assumption)
	(a)	$\mathsf{Tr}_c(p)$		(assumption)
		(a.i)	λ_c	(1., (a), eq. (2.1))
		(a.ii)	$\neg\exists_c p(\mathsf{Ext}(\lfloor\lambda_c\rfloor, p, c) \wedge \mathsf{Tr}_c(p))$	(def. λ_c)
		(a.iii)	$\neg\mathsf{Tr}_c(p)$	(1., (a.ii), eq. (2.3))
		(a.iv)	$\mathsf{Tr}_c(p) \wedge \neg\mathsf{Tr}_c(p)$	((a), (a.iii))
	(b)	$\neg\mathsf{Tr}_c(p)$		(assumption)
		(b.i)	$\neg\lambda_c$	(1., (b), eq. (2.1))
		(b.ii)	$\exists_c p(\mathsf{Exp}(\lfloor\lambda_c\rfloor, p, c) \wedge \mathsf{Tr}_c(p))$	((b.i), def λ_c)
		(b.iii)	$\mathsf{Tr}_c(p)$	(1., (b.ii), eq. (2.3))
		(b.iv)	$\mathsf{Tr}_c(p) \wedge \neg\mathsf{Tr}_c(p)$	((b), (b.iii))
2.	$\neg\exists_c p(\mathsf{Exp}(\lfloor\lambda_c\rfloor, p, c))$			(1. – (b.iv))
3.	$\neg\exists_c p(\mathsf{Exp}(\lfloor\lambda_c\rfloor, p, c) \wedge \mathsf{Tr}_c(p))$			(2., logic)
4.	λ_c			(3., def. λ_c)
5.	$\exists_c p(\lfloor\lambda_c\rfloor, p, c)$			(eq. (2.2))
6.	$\exists_c p(\lfloor\lambda_c\rfloor, p, c) \wedge \neg\exists_c p(\lfloor\lambda_c\rfloor, p, c)$			(2., 5.)

The key insight of contextualism is that there is a context shift between steps 4. and 5. In other words, the Liar argument establishes that λ_c does not express a true proposition in c (4.). But then, context shifts to a subsequent context c' and instead of 5. we have

5′. $\exists_{c'} p(\lfloor\lambda_c\rfloor, p, c')$ (eq. (2.2), context shift)

which blocks the contradiction. The intuition behind this approach is that we can reason in a given context c until we reach the paradoxical λ_c, which says, according to the contextualist interpretation, that λ_c does not express a true proposition in c. But if we reach λ_c, then by eq. (2.2), λ_c expresses a proposition after all. However, this insight takes us to a subsequent context c' so that now, in c', we can reason that λ_c expresses a true proposition in c', namely the proposition that λ_c does not express a true proposition in c.[30]

Glanzberg develops his contextualism building on Kripke's seminal paper "Outline of a Theory of Truth." Kripke's main idea is a model-theoretic construction of an interpretation of the truth predicate based on the three-valued strong

30 In order to rephrase the contextualized Liar argument without propositions, we have to drop the Exp-predicate. Instead, we define λ_c simply as $\neg\mathsf{Tr}_c(\lfloor\lambda_c\rfloor)$. The assumption that λ_c expresses a true proposition in c can be rephrased as $\mathsf{Tr}_c(\lfloor\lambda_c\rfloor) \vee \neg\mathsf{Tr}_c(\lfloor\lambda_c\rfloor)$. Then, the argument follows a similar pattern. See Rossi [144, sec. 2] for details.

Kleene evaluation scheme.[31] The construction yields a pair $\langle E, A\rangle$, where E is the extension and A is the antiextension of the truth predicate, corresponding to the truth values 1 and 0 respectively. Since I will only need E in what follows, I mostly focus on the construction of E.

The construction of E proceeds in stages indexed by ordinals, so that E is the limit of a sequence $E_0, E_1, ..., E_\alpha, ...$[32] Start with a model of the truth-free fragment of the language.[33] At stage 0, E_0 is empty. Then, all truth-free atomic base sentences (such as $0 = 0$) which are verified by the arithmetic base structure are added to E_1. Furthermore, all negations of truth-free atomic base sentences (such as $\neg(0 \neq 0)$) that are falsified by the arithmetic base structure are also added to E_1. We go from E_1 to E_2 by applying the strong Kleene evaluation schemes. For example, $\varphi \wedge \psi$ is in E_2 if and only if both, φ and ψ are in E_1, and $\neg\neg\varphi$ is in E_2 if and only if φ is in E_1. Furthermore, E_2 contains $\mathsf{Tr}(\lfloor\varphi\rfloor)$ for all sentences φ in E_1, and all sentences $\neg\mathsf{Tr}(\lfloor\varphi\rfloor)$, for all sentences $\neg\varphi$ in E_1. All subsequent stages are constructed in a similar way, i.e., by applying the truth predicate and closing under strong Kleene evaluation. For cardinality reasons, the resulting sequence $E_0, E_1, ..., E_\alpha, ...$ reaches a stage where no new sentences are added, i.e., a stage such that $E_{\alpha+1} = E_\alpha$. Such a stage is called a *fixed point*. A fixed point contains all true sentences, not only of the truth-free part of the language but of the whole language. A fixed point is *minimal* if it is contained in all other fixed-points that follow the same construction rule, and *nonminimal* otherwise. In the case of the Kripke construction, we get the minimal fixed point when we start with $E_0 = \emptyset$, and a non-minimal fixed point when $E_0 \neq \emptyset$. A fixed point is *consistent* when there is no sentence φ such that φ and $\neg\varphi$ are in E. The construction of the antiextension runs parallel to the construction of the extension, but instead of taking the verified sentences, one takes the falsified sentences.[34] Alternatively, one can always retain the antiextension from the extension as the set of all sentences $\neg\varphi$ for φ a sentence from the extension.

Essential to the picture is that the Kripke construction proceeds bottom-up, i.e., we start with atomic base sentence and construct more and more complex sentences via logical operations and application of the truth predicate. Kripke calls such sentences *grounded*, whereas sentences such as the Liar are *ungrounded* and never considered during the construction as they are circular by definition.[35] The truth predicate is therefore a partial predicate, i.e., the union of the extension and the antiextension is only a proper subset of the sentences of the language. Sentences such as the Liar that are neither in E nor in A are in the *gap*.

31 The strong Kleene evaluation scheme adds a third truth value, 1/2, which can be interpreted as 'neither true nor false'. For more details on the strong Kleene evaluation scheme and the Kripkean construction, see McGee [102, ch. 4], Field [34, ch. 3], and Halbach [57, ch. 15].

32 For the formal details on fixed point constructions and inductive definitions, see section 3.2.

33 Typically, the truth-free fragment of the language is just the language of arithmetic, so think of the standard model of arithmetic $\mathbb{N}$.

34 Since Kripke's construction uses three-valued logic, the falsified sentences are not just those sentences that are not verified, as this would give the collection of sentences with truth value 1/2 and 0.

35 The philosophical intuition is that grounded sentences are somehow *grounded in facts*, whereas ungrounded sentences never bottom-out in claims about arithmetic base facts but contain self-referential loops. See Rossi [143] for details.

The important step for the contextualist now is that there is an open-ended sequence of contexts, and every context needs its own extension of the truth predicate. Moreover, recall that we want to have the logic to be classical, so the standard Kripke construction has to be adjusted. In order to have classicality, Glanzberg takes the *closing-off* of the minimal fixed point for the initial context c, i.e., the set of sentences φ such that $\langle \mathcal{M}, E \rangle \models \varphi$. Most important is that, since λ_c is not in E, we have that $\langle \mathcal{M}, E \rangle \not\models \mathsf{Tr}_c(\lfloor \lambda_c \rfloor)$. Since we use the classical consequence relation $\models$, we get $\langle \mathcal{M}, E \rangle \models \neg\mathsf{Tr}_c(\lfloor \lambda_c \rfloor)$ and so, by definition of λ_c, $\langle \mathcal{M}, E \rangle \models \lambda_c$. This means that λ_c is in the closing-off, $\mathsf{ClOff}(E)$, of the minimal fixed point.

In order to model the context shift, we run the Kripke construction again for a new truth predicate $\mathsf{Tr}_{c'}$, but instead starting at stage 0 from $\emptyset$, we start with the closing-off of E, $\mathsf{ClOff}(E)$, i.e., the sentences φ such that $\langle \mathcal{M}, E \rangle \models \varphi$. The new Kripke construction yields another extension, E' for $\mathsf{Tr}_{c'}$ which contains λ_c, and we can again take the closing-off for context c', $\mathsf{ClOff}(E')$. Now, since λ_c is in the closing-off of context c' $\mathsf{ClOff}(E)$, we get that $\langle \mathcal{M}, \mathsf{ClOff}(E') \rangle \models \lambda_c$, but again, $\langle \mathcal{M}, \mathsf{ClOff}(E') \rangle \not\models \lambda_{c'}$. The result yields exactly the described open-ended sequence of non-minimal Kripke fixed points where each step corresponds to a context.[36]

The contextualist solution to the Liar paradox is hierarchical. Irrespective of whether we look at the contextualist interpretation of the Liar paradox from a proof-theoretical or from a model-theoretical perspective, the picture is always the same: contextualist postulate an open-ended (possibly transfinite) sequence of contexts, where each context essentially builds on (all) previous contexts. The order of these contexts induces a hierarchy that we can index by ordinals $c_0, c_1, ..., c_n, ..., c_\alpha, ...$ This is most obvious in Glanzberg's model-theoretic construction, where the hierarchy is given by set-theoretical inclusion. For every context $c_\alpha, c_{\alpha+1}$, the closing-offs are ordered by the proper subsets relation, i.e. $\mathsf{ClOff}(E_\alpha) \subseteq \mathsf{ClOff}(E_{\alpha+1})$. Hence, every context *strictly expands* all previous contexts, so that we get a hierarchy of context-depending interpretations of the truth predicate.

2.2.3 *Williamson's Version of Russell's Paradox*

Let us finally consider a paradox due to Timothy Williamson [169] which is an application of the pattern of reasoning behind Russell's paradox to a concept that is much broader in scope than the set-concept. The concept in question is the concept of an interpretation of a (first-order) object language, be it the language of arithmetic, truth, set theory or just natural language.

The paradoxes considered so far had disruptive consequences for the naïve concepts they proved inconsistent. In the case of truth, it was the requirement that the language is powerful enough to encode its own syntax to form self-referential sentences, and that the truth predicate satisfies the *naïveté* principles; in the case of set theory, it was the intuition that sets are extensions of predicates.

36 The material presented in this paragraph is a deviation from Glanzberg's original work in Glanzberg [47] due to Rossi [144]. In the original presentation, Glanzberg works only with a single truth predicate, but this requires much more technical effort.

Williamson's merit is that he was able to generalize the paradoxical reasoning so as to affect already a naïve conception of interpretation. This implies that any kind of naïve semantic theorizing is prone to paradoxical reasoning. Consider an object language $\mathcal{L}$ and assume that we are free to interpret any predicate symbol with any metalanguage property:

> For any relation symbol R and property P, there is an interpretation I_P that interprets R as P. (2.5)

As such, Williamson's version has consequences that are as disruptive for a naïve understanding of the concept *interpretation* as Russell's paradox and the Liar paradox are for the naïve principles of comprehension and truth. And just like the solutions developed in the previous sections, the solution to Williamson's version of Russell's paradox will be hierarchical. As in the case of Russell's paradox, the ability to quantify over absolutely everything will play an important role in Williamson's version of Russell's paradox as one straightforward reaction is to reject this assumption. One further assumption is that we can speak and quantify over interpretations just as we can speak and quantify over any other object:

> Interpretations are objects. (2.6)

Here is Williamson's version of Russell's paradox:[37] Let $\mathcal{L}$ be a first-order object language, R be a unary predicate symbol, D be the domain of quantification, and P be a metalanguage predicate. Then, there should be an $\mathcal{L}$-interpretation that interprets R with P, call this interpretation I_P. More precisely, I_P is such that

1. $\forall x$(I_P is an interpretation according to which 'R' applies to $x \leftrightarrow Px$).

Now consider the following predicate Q:

2. $\forall x$($Qx \leftrightarrow x$ is not an interpretation according to which 'R' applies to x).

By 2.5, we're free to substitute Q for P, so that we get

3. $\forall x$(I_P is an interpretation according to which 'R' applies to $x \leftrightarrow$ x is not an interpretation according to which 'R' applies to x).

Finally, by 2.6, we can quantify over I_P just as we can quantify over any object whatsoever, i.e., if I_P is in D, we can eliminate $\forall x$ and instantiate it by I_P, so we get

4. I_P is an interpretation according to which 'R' applies to $I_P \leftrightarrow$ I_P is not an interpretation according to which 'R' applies to I_P.

37 My presentation follows Linnebo [89].

Just like in the case of Russell's paradox, one straightforward reaction is to argue that I_P is not in D. Note that this is not a rejection of 2.6, but only means that D is restricted in such a way that it does not contain I_P. If I_P is not in D, the contradiction is blocked because the last step from 3. to 4. cannot be done. If I_P is not in D, we cannot instantiate x in the $\forall$-elimination step with I_P, and so we don't arrive at the paradoxical conclusion. Now assume that D contains absolutely everything. Then, if we keep assumption 2.6, the paradoxical conclusion cannot be blocked. If D contains absolutely every*thing*, and if interpretations are objects (and so *a fortiori* things), I_P must be in D and the instantiation of x with I_P in the final step is legitimate.

At this point, we have two possible assumptions to reject. One is assumption 2.6 that interpretations are objects, and the other is absolute generality. Since the interest of this work lies in defending absolute generality to some degree, let us consider the reaction to the paradox that is based on the rejection of assumption 2.6.

The rough idea is to think of interpretations not as objects, but instead to use higher-order logic to formalize interpretations as relations between expressions of the language and the world.[38] For simplicity, I will from now on speak about models and not about interpretations as this makes the presentation easier. Let's start from the standard definition of a model. Let $\mathcal{L}$ be a first-order object language as above. In standard model-theory, a model for $\mathcal{L}$ is defined to be an ordered pair $\mathcal{M} = \langle D, I \rangle$, where D is a non-empty set—the domain of the interpretation—and I is a function—coded as a set—that assigns semantic values to non-logical expressions of $\mathcal{L}$. As such, $\mathcal{M}$ is a set, and therefore an object.

The alternative due to Rayo and Uzquiano [135] considers a model not in terms of sets, but in terms of higher-order logic.[39] The notion of a model $\mathbb{M}$ for a first-order language—I use $\mathbb{M}$ for higher-order models and $\mathcal{M}$ for model-theoretic models—is given by a second-order formula that defines what it takes for $\mathbb{M}$ to be an interpretation of the object language. The second-order formula will have a single free unary second-order variable so that the picture will be roughly as follows:

> $\mathbb{M}(X)$ if and only if X satisfies the higher-order *model-predicate*.

$\mathbb{M}$ specifies that X contains a higher-order analogue of the domain which satisfies the non-emptiness requirement, that X contains an interpretation for every non-logical expression of the language, and so on. So instead of saying that $\mathbb{M}(X)$ is a model, it would be more accurate to say that $\mathbb{M}(X)$ defines the *model-relation*, which contains clauses specifying that for an $\mathcal{L}$-constant c, and an object o, c stands in the model-relation to o, i.e., according to $\mathbb{M}(X)$, the interpretation of c is o, and so on.[40]

38 This approach goes back to Boolos [12]. The view has been discussed and received further endorsement by Cartwright [21], Williamson [169], Linnebo [89], Rossi [144], Florio and Jones [40, 39], Lewis [83], and Trueman [158]. It has also been developed in technical detail in Rayo and Uzquiano [135], Rayo [132], Linnebo and Rayo [93], Rayo and Williamson [136], and Button and Walsh [18].

39 A full presentation of the alternative account is given in section 3.4.

40 Even though it would be more accurate to avoid the reification of the *model-relation* given by $\mathbb{M}(X)$, I will often fall back to 'object-talk' for simplicity.

In the context of absolute generality, the interpretation of the second-order variable X is of great importance. According to the standard interpretation, second-order expressions denote sets. Interpreted that way, X is an object, which is exactly what we wanted to avoid. Instead of the standard interpretation, we use the *plural interpretation* that has been developed by George Boolos in the 1980's. Roughly, the idea is to introduce a new kind of denotation—*plural denotation*—according to which X does not singularly denote a set-theoretic object, but rather denotes many objects *simultaneously*. The plural interpretation of the second-order variables does not commit us to entities beyond those contained in the first-order domain. This is called the *ontological innocence of plural quantification*.[41]

Coming back to Williamson's version of Russell's paradox, the strategy is to claim that an interpretation of a first-order object language is given by the values of a unary second-order variable (i.e. interpreted as an ontologically innocent plurality), and therefore the interpretation I_P that caused the troubles in step 4. is not in the relevant quantifier domain. Moreover, since interpretations are developed in terms of pluralities, they are not considered to be objects at all, and so they do not commit us to reified relations. However, now that we have developed the notion of an interpretation of a first-order object language in a second-order metalanguage, the question emerges what about the second-order metalanguage? As Williamson [169] observes, this interpretation can only be given in a metalanguage of yet a higher order, a metametalanguage that is a third-order language. At this point, we have to operate essentially as in the case of the interpretation of a first-order object language and develop interpretations of second-order languages as unary third-order predicates that are interpreted with a *super plurality*, i.e. a *plurality of pluralities*. It should now be obvious that the solution to Williamson's version of Russell's paradox is hierarchical: the above line of reasoning leads naturally to an open-ended sequence of higher-order metalanguages where every object language of order n is interpreted by a metalanguage of order $n + 1$.[42]

Let me summarize: In this section, I have discussed two of the most important paradoxes in the history of logic, mathematics, and philosophy—Russell's paradox and the Liar paradox—as well as a version of Russell's paradox due to Timothy Williamson. I have argued that the most promising solutions to all these paradoxes share a common feature, and that is their *hierarchical* nature. In the next section, I will show how relativists use the paradoxes to argue against absolutism. I will come back to the hierarchical nature of the solutions in section 2.4.

41 Chapter 3 contains background material to all the topics mentioned here, in particular on higher-order logic (section 3.1.2) and the plural interpretation (section 3.1.3). A throughout discussion on the interpretation of higher-order expressions in the context of absolute generality is given in section 6.1.

42 See Rayo [132] and Linnebo and Rayo [93] and section 3.4 for details.

2.3 RELATIVISM AND PARADOXES

By far the strongest and most frequently used argument for relativism is based on paradoxes.[43] Key to all these arguments is that the paradoxes reveal a certain property of the underlying conception that is incompatible with absolutism. This property is sometimes called a *reproductive process* [146], *indefinite extensibility* [31], or *diagonalization* [150].

Within this section, I will first show how this three phenomena are used as explanations for the paradoxes (section 2.3.1), and then how relativists use this to argue against absolutism (section 2.3.2).

2.3.1 *Reproductive Processes, Indefinite Extensibility, and Diagonalization*

Russel identifies self-reproductive processes as the source of paradoxes, and describes them as follows:[44]

> [T]he contradictions result from the fact that, according to current logical assumptions, there are what we may call self-reproductive processes and classes. That is, there are some properties such that, given any class of terms all having such a property, we can always define a new term also having the property in question. Hence we can never collect all the terms having the said property into a whole; because, whenever we hope we have them all, the collection which we have immediately proceeds to generate a new term also having the said property. [146, p. 36]

In other words, the self-reproductive nature of some concepts C (e.g., 'set') allows us to generate, from any collection W of objects falling under C, a new object which also falls under C but is not in W.

A successor of Russell's notion of reproductive processes is given by Michael Dummett, who has famously argued that the *indefinite extensibility* of the concept 'set' shows that quantification over all sets is impossible. Dummett characterizes indefinite extensibility as follows:

> An indefinitely extensible concept is one such that, if we can form a definite conception of a totality all of whose members fall under

43 See Linnebo [89], Rayo [132], Weir [166], Williamson [169], McGee [101], Shapiro [148], Fine [36], Glanzberg [48], Parsons [119], Florio [38], Cartwright [21], Dummett [31, 29], Rossi [144], and Studd [152] which all consider the paradoxes to be the strongest argument against absolutism.

44 In the following, I often speak of "diagonalization," "indefinite extensibility," and "self-reproductive processes" in closely related terms, sometimes even interchangeably. This reflects a structural similarity: certain totalities, when treated as completed totalities, can be used to generate new elements not already included within them. In that sense, the diagnoseses that concepts are *indefinitely extensible*, lead to *self-reproductive processes*, or that it is possible to *diagonalize out of certain totalities* all suggest a kind of open-endedness of the respective concept. However, I do not mean to suggest that these notions are strictly identical. Diagonal arguments, for instance, have a precise formal characterization (as in Cantor's theorem or Gödel's Incompleteness theorem), while a formal characterization of indefinite extensibility, especially in its Dummettian formulation, is much more difficult. For a more nuanced treatment of the relationships between Russell's *self-reproductive processes* and Dummett's *indefinite extensibility*, see Linnebo [85]. Fore a more detailed of the relation between *indefinite extensibility* and *diagonalization*, see Simmons [150]. See also section 6.2.4.

> that concept, we can, by reference to that totality, characterize a larger totality all of whose members fall under it. [31, p. 441]

Just like Russell's self-reproductive processes, Dummet characterizes a concept C as indefinitely extensible if, for any collections W of objects falling under C, it is possible to construct an object c that also falls under C but is not contained in W, and all that is needed to construct c is to refer to W.

In a more recent publication, Luca Incurvati [73] defines a concept C to be indefinitely extensible if and only if "whenever we succeed in defining a set u of objects falling under C, there is an operation which, given u, produces an object falling under C but not belonging to u" [73, p. 27]. The notion of a self-reproductive process or an indefinitely extensible concept leads to a contradiction when we combine it with a principle of universality: "A concept C is universal iff there exists a set of all the things falling under C" [ibid.]. According to Dummett, the concept 'set' is indefinitely extensible:

> Russell's concept of a [set] not containing itself as a member is a prototypical example of an indefinitely extensible concept: for, once we form a definite conception of a totality R of such [sets], it is evident that R cannot, on pain of contradiction, be a member of itself, and thus the totality consisting of all the members of R together with R itself, is a more extensive totality than W of [sets] that are not members of themselves. [30, p. 317, notation adjusted]

So Russell's paradox implies that the concept 'set' is indefinitely extensible and when combined with the principle of universality, yields a paradox: for every set of sets K, by Russell's paradox, there exists a set outside of K. Now assume the principle of universality for sets, i.e., assume that there exists a set of all sets U. We get a contradiction because the indefinite extensibility of the set-concept generates a set that is not contained in the universal set.

Similarly the Liar paradox shows the indefinitely extensibility of the concept 'proposition' in the contextualist solutions: take any collection of proposition for the initial context c. With some basic assumptions about the underlying language, we can form a Liar sentence λ_c that forces us to step outside the initial context c into a subsequent context c' which contains a proposition corresponding to λ_c. If we assume the principle of universality for propositions, i.e., that there is a context which contains a collection of all propositions, we get a contradiction because the indefinite extensibility of the concept 'proposition' generates a proposition not contained in the collection of all propositions.[45]

Another line of reasoning that leads essentially to the same conclusion is given by Keith Simmons [150], who develops his views based on the notion of *diagonal arguments*. Cantor was the first to use diagonal argument to prove

45 In order to reformulate this without using propositions, we can rephrase the above reasoning and use sentences instead of propositions as truth-bearers. For any initial context c, the assumption that the Liar sentence is either true or not true in c ($\mathsf{Tr}_c(\lfloor\lambda_c\rfloor) \vee \neg\mathsf{Tr}_c(\lfloor\lambda_c\rfloor)$) leads to a contradiction and therefore causes a context shift to a subsequent context c' in which we can determine the truth value of λ_c. In a proposition-free approach to contextualism, it is the the generality of the truth predicate that expands. Just like proposition-based contextualism, the assumption of universality—i.e. the assumption of a maximally general truth predicate that is not prone to expansion—leads to an inconsistency.

the existence of uncountable sets: consider the set of infinite binary sequences $B = \{\langle b_0, b_1, ...\rangle \mid b_i = 0 \text{ or } 1\}$. Assume that B is countable, i.e. there exists an enumeration such that each $b^0, b^1, b^2, ...$ is an infinite binary sequence. Then, we can write down all elements in the series as an array with the following form:

$$\begin{array}{lccccccc}
b^0 = & b_0^0 & b_1^0 & b_2^0 & \dots & b_3^0 & \dots \\
b^1 = & b_0^1 & b_1^1 & b_2^1 & \dots & b_3^1 & \dots \\
b^2 = & b_0^2 & b_1^2 & b_2^2 & \dots & b_3^2 & \dots \\
b^3 = & b_0^3 & b_1^3 & b_2^3 & \dots & b_3^3 & \dots \\
\vdots & \vdots & \vdots & \vdots & \dots & \vdots & \ddots
\end{array}$$

where each b_j^i is either 0 or 1. Now we take the diagonal sequence $c = \langle b_i^i \rangle$ and form the antidiagonal by flipping each digit. This yields the antidiagonal sequence $d = \langle 1 - b_i^i \rangle$ which is 0 if b_i^i is 1 and which is 1 if b_i^i is 0. By definition, d is not in the above list because for every sequence b^j, d and b^j differ at least at position j. Simmons argues that Russell's paradox and the Liar paradox can be understood as diagonal arguments. Consider Russell's paradox: assume that the sequence x_α (for α an ordinal) lists every set. We get an array by considering the following cases:

$$R(x, y) = \begin{cases} 1, & \text{if } y \in x, \\ 0, & \text{if } y \notin x \end{cases}$$

i.e., if the resulting array shows 1, then the leftmost set is an element at top of the row:

	x_1	x_2	x_3	x_4	$\dots$
x_1	x_1^1	x_1^2	x_1^3	x_1^4	$\dots$
x_2	x_2^1	x_2^2	x_2^3	x_2^4	$\dots$
x_3	x_3^1	x_3^2	x_3^3	x_3^4	$\dots$
x_4	x_4^1	x_4^2	x_4^3	x_4^4	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

where each x_j^i is either 0 or 1. The diagonal corresponds to the set of all self-membered sets. In the iterative picture, the diagonal sequence contains only 0. But here, we assume the naïve conception. Assume that there is a Russell set of all non-self-membered sets which appears at position α in the list. The Russell set would correspond to an antidiagonal $R = \langle r_\alpha \rangle$ given by

$$r_\alpha = \begin{cases} 1, & \text{if } x_\alpha^\alpha = 0, \\ 0, & \text{if } x_\alpha^\alpha = 1. \end{cases}$$

But this antidiagonal cannot appear in the array, as the following array shows:

	x_1	x_2	x_3	x_4	…	R	…
x_1	x_1^1	x_1^2	x_1^3	x_1^4	…	$1 - x_1^1$	…
x_2	x_2^1	x_2^2	x_2^3	x_2^4	…	$1 - x_2^2$	…
x_3	x_3^1	x_3^2	x_3^3	x_3^4	…	$1 - x_3^3$	…
x_4	x_4^1	x_4^2	x_4^3	x_4^4	…	$1 - x_4^4$	…
⋮	⋮	⋮	⋮	⋮	⋱	⋮	…
R	x_α^1	x_α^2	x_α^3	x_α^4	…	???	…
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋱

The construction of the antidiagonal has not enough information to determine every cell. The one containing ??? is the cell corresponding to the question whether $R \in R$. The paradox arises in the diagonal argument as the fact that some cells cannot be determined.

Just as with indefinite extensibility, the procedure of diagonalization can be used to explain the paradoxes. Consider the Liar paradox: in the initial context c, we form the Liar sentence λ_c which diagonalizes out of c into the subsequent context c' that contains the proposition corresponding to λ_c. Since we can diagonalize out of any context, no context can contain all propositions. Similarly with Russell's paradox: constructing the antidiagonal as above leads to a contradiction. We diagonalize out of any allegedly all-inclusive collection of sets.

It is important to note that the notions of *indefinite extensibility* and of *diagonalization* essentially rely on the fact that the solutions to paradoxes are hierarchical: when we look at any stage in the cumulative hierarchy V_κ, considering it's Russell set R_κ extends or diagonalizes out or V_κ and essentially pushes us to the next stage in the hierarchy which properly extends the underlying stage. Similarly with respect to the Liar paradox: contextualism interprets the argument from the Liar sentence as a way to extend or diagonalize out of the current context and essentially pushes us to the next stage in the hierarchy of contexts, thereby switching to a subsequent context that properly extends the previous context. So the hierarchical nature of the solution to the paradoxes and the indefinite extensibility (or the suitability for diagonalization) are just two sides of the same coin. Let me now show how relativists use the paradoxes to argue against absolutism.

2.3.2 *The Relativistic Argument from Paradox*

The fundamental assumption of absolutism is that there is a domain, call it M, that contains absolutely everything. At this point, it is important to ask which kind of object M is? A strong candidate certainly is to assume that M is a set, but other options are also possible. Let us, for the moment, just assume that M is a *collection*. I thereby mean that M is an *entity* that contains other objects, that M is an entity different from all its members, that M is non-empty (at least *something* exists), and that M satisfies some form of the extensionality principle (i.e. if N is any collection that contains exactly the same elements that M contains, then $M = N$). Just like the concept 'set', the so-described concept

'collection' is obviously indefinitely extensible or prone to diagonalization: if R_M is the subcollection of all non-self-membered collections in M, than R_M cannot be in M.

Having established the indefinite extensibility or the suitability to diagonalization of the concept 'collection', let me now give the *relativistic argument from paradox*: Assume that M is the domain of the absolutist, i.e., M is the domain of the quantifier that, as the absolutist claims, ranges over absolutely everything. By reasoning along the lines of Russell's paradox, this leads into the paradoxical conclusion that R_M is in R_M if and only if R_M is not in R_M. In order to avoid this inconsistency, we *extend the domain* or *diagonalize out of the domain*, and claim that R_M is not in M, i.e., we have identified an object that is not contained in the allegedly absolutists domain. This implies that there is a larger domain M' containing at least M and R_M. Hence, M cannot be the domain in the sense of absolutism as M does not contain absolutely everything.

We can reason analogously in the case of to the Liar paradox: if M is the domain containing absolutely everything, it contains, among many other things, absolutely every proposition. By reasoning along the line of the Liar paradox, this leads into the paradoxical situation that the proposition corresponding to the Liar sentence is true if and only if it is not true. In order to avoid this inconsistency, we adopt the contextualist method and claim that in the initial domain, there is no proposition corresponding to the Liar sentence. We *extend the domain* or *diagonalize out of the domain*, and find the proposition corresponding to the Liar sentence outside the allegedly absolutist domain. This implies that there is a larger domain M' containing at least M and the proposition corresponding to the Liar paradox. Hence, M cannot be the domain in the sense of absolutism as M does not contain absolutely every proposition. In both cases, we can repeat the reasoning, this time taking M' as the starting domain, to reach yet a wider domain M'', and so on. Since M was arbitrary, the relativistic argument from paradox disqualifies any allegedly all-inclusive collection from being the domain in the sense of generality absolutism.

The relativistic argument from paradox rule out any approach to semantics that defines domains to be collection-like entities from being compatible with absolutism. In particular, it disqualifies standard model theory—which defines domains to be sets—from modeling unrestricted quantification. Michael Dummett concludes that "it is, [...] the prime lesson of the discovery of the set-theoretical paradoxes that quantification over" a domain containing "the totality of all sets" or "the totality of all objects whatsoever" is impossible [29, p. 516]. Similarly, Steward Shapiro "take[s] it as agreed that the antinomies provide the basic motivation for generality relativism" [148, p. 469]. And also Michael Glanzberg summarizes that regarding an absolutely unrestricted quantifier, we "conclude on pain of contradiction that [some] object cannot be in the domain of the quantifier [and] that our apparently absolutely unrestricted quantifier ... is somehow restricted after all" [48, p. 47]. In other words, relativists interpret the argument from paradox as proving absolutism wrong.

Note that the relativistic argument from paradox relies on the hierarchical nature of the solutions to the paradoxes. Whenever we extend or diagonalize out of an allegedly all-inclusive collection, according to relativism, the reason-

ing pushes us to the next stage in the hierarchy. This means that we either step to the next level of the cumulative hierarchy which contains all previous levels and all Russell sets of all previous levels, or into the next context which contains all previous context and all propositions corresponding to all previous Liar sentences. However, this level or context, again, is prone to extensibility or diagonalization and we're pushed again to the next layer or context. Since no layer or context is immune to diagonalization or extensibility, no layer or context can be universal.

Let me summarize: in this section, I have argued that relativists identify *indefinite extensibility* and *diagonalization* as the main sources of the paradox. However, only when combined with the assumption of universality—i.e., the assumption that there is an all-inclusive collection containing all instances of a particular concept—these notions really cause an inconsistency. Relativists interpret this fact as showing that absolutism must be wrong and, according to the relativistic picture, no domain can achieve absolute generality. Moreover, I have argued that *diagonalization* and *indefinite extensibility* on the one hand, and the hierarchical nature of the solution to the paradoxes on the other hand, are just different ways to think about the matter. In the next section, I will show how to combine the hierarchical solutions to the paradoxes with absolute generality to get the best of two worlds.

2.4 THE BEST OF TWO WORLDS

2.4.1 *Reassessing the Relativistic Argument from Paradox*

We saw that relativists interpret the paradoxes as proving absolutism wrong. But have they really succeeded in establishing this conclusion? Recall that the strongest argument against absolutism is the argument from paradox, but what does it show? Relativists start from an allegedly all-inclusive collection-like entity M which they take to be the domain of the unrestricted quantifier, and then argue from paradox: if M is all-inclusive, it contains its Russell-collection if and only if it does not contain it, hence the paradox.

This shows that the interpretation of the sentence postulating the Russell-collection of M leads to an inconsistency when its quantifiers range over absolutely everything. More precisely, if in $\exists y \forall x (x \in y \leftrightarrow x \in M \wedge x \notin x)$ the quantifiers range over absolutely everything, we get in troubles: we instantiate x and y with R_M and get $R_M \in R_M \leftrightarrow R_M \in M \wedge R_M \notin R_M$. Since M is all-inclusive, $R_M \in M$ is vacuously true, and so we get $R_M \in R_M \leftrightarrow R_M \notin R_M$.

This certainly shows that paradoxes prove absolutism wrong for sentences crucial for paradoxical reasoning. Let me explain: recall from section 2.2 and section 2.3 that paradoxes rely on the indefinite extensibility or the suitability for diagonalization of the concepts 'set' or 'proposition'. In other words, the expansion of the universe of discourse is connected to paradoxes in the sense that sentences causing the paradoxes are the driving force behind domain expansion. So, some sentences—call them *expansion sentences*—trigger the expansion procedure. The main lesson of the forgoing sections is that expansion sentences

must not be interpreted over an all-inclusive domain to avoid paradoxes. This is most obvious in the contextualist version of the Liar argument. Recall from section 2.2.2 that in the following step:

4.	λ_c	(3., def. λ_c)
5′.	$\exists_{c'} p(\lfloor \lambda_c \rfloor, p, c')$	(eq. (2.2), context shift)

we shifted from the initial context c to the subsequent context c'.[46] However, the fact that it shifts is the main ingredient of the contextualist solution and in the case of the Liar paradox, it has certainly to do with the expressive resources that are enlarged in light of the Liar sentence. In the contextualist solution, this is implemented by domain expansion from a starting domain not containing a proposition corresponding to the initial Liar sentence, to a subsequent domain containing a proposition corresponding to the initial Liar sentence.[47] Similarly for Russell's paradox, where the domain expansion gets triggered by the respective Russell sentence, and causes us to step up in the cumulative hierarchy.

The relativistic argument from paradox shows that expansion sentences have to be subject to domain restrictions in order to avoid paradoxes. But the prima facie examples from absolutists typically are not expansion sentences. Instead, they invoke prima facie absolutistic examples such as $\forall x\ x = x$, or $\neg\exists x\ x \in \emptyset$. None of these examples, however, are expansion sentence. There is no context shift or otherwise domain-expansion mechanism involved in the sentence $\forall x\ x = x$. It therefore follows that the interpretation of the relativistic argument from paradox—that the argument proves absolutism wrong—is too strong. The argument does not refute absolutism *tout court*, it only refutes absolutism for expansion sentences. This sets the stage to describe *bicontextualism* in the next section as an intermediate position between absolutism and relativism.

2.4.2 *Bicontextualism*

I have described the cumulative-iterative picture of modern set theory, as well as truth-theoretic contextualism, which are among the most elegant and powerful solutions to the paradoxes. As we saw, however, they are incompatible with absolutism—not with absolutism in general, but with absolutism for expansion sentences. On the other hand, as we saw in section 2.1, absolutism has some nice features, none of which relies on absolutism for expansion sentences. It is the purpose of this section to describe how we can, and why we should, put these two features together and develop an intermediate position called *bicontextualism* [144] that combines all the nice features of absolutism (for non-

46 How this process of domain expansion exactly proceeds is an open question and beyond the scope of this work. Some accounts are on offer: Glanzberg [48], Gauker [46], Fine [36, 35, 37], and Mankowitz [99], Studd [152, chs. 4.4, 8].

47 As mentioned before, propositions can always be paraphrased away. If we take sentences as the primary truth bearers, the contextualized liar sentence λ_c leads to a contradiction that causes a context shift where a new truth predicate can be used to express the fact that λ_c is *true-in-c′*. But then, the interpretation of Tr_c was not maximally general. This way, indefinite extensibility of diagonalization affects the proposition-free version of the liar just as well. See also footnote 45.

expansion sentences) with the elegant and powerful hierarchical solutions of the paradoxes.

Consider the language of set theory $\mathcal{L}_\in$. As Rossi [144, sec. 4] points out, in model theory we usually apply a unified interpretation to all $\mathcal{L}_\in$-sentences in the sense that we fix a model $\mathcal{M} = \langle M, I \rangle$ and define $\models$ as a relation between $\mathcal{M}$ and the $\mathcal{L}_\in$-sentences. This means that the quantifiers of all $\mathcal{L}_\in$-sentences range over the domain M. The idea now is to drop this uniformity-approach to interpretations and instead allow, in the interpretations, that the quantifier domains vary. First, assume that we have a non-objectual, all-inclusive plural 'domain' containing absolutely everything.[48] Since this 'domain' contains absolutely everything, it contains absolutely every set, and hence, every rank of the cumulative hierarchy V_κ. Once we've dropped the uniformity-approach, we can interpret some of the $\mathcal{L}_\in$-sentences with the all-inclusive 'domain', and others only with restricted domains V_κ. The important question now is which sentences we should interpret over which domain, but in light of the discussion of the previous sections, there is already a hint how to proceed: let's take those sentences that cause domain expansion in some way, i.e., in the case of set theory, most notably expansion sentences as involved in Russell's paradox, and interpret them only over the restricted domains V_κ. Sentences that are not expansion sentences, on the other hand, are unproblematic, and we can freely interpret them over the non-objectual all-inclusive plural 'domain'. In a slogan, bicontextualism is based on the following idea:

> **Bicontextualism:** Allow absolutism for all non-expansion sentences and enforce relativism for all expansion sentences.

According to bicontextualism, we ultimately want to divide the sentences of $\mathcal{L}_\in$ into those sentences that we can interpret over the all-inclusive 'domain', and into those sentence that we can interpret only over restricted domains. Call the former the *inherently absolutistic sentences* (Abs), and the later the *inherently relativistic sentences* (Rel). It is important not to presuppose the distinction between Abs and Rel, as this would weaken the position of bicontextualism significantly. Rather, the distinction comes out from the formal theory. Let me briefly sketch how this is done for the case of set theory:

Let V be the universe of sets.[49] As already noted, V contains all nice standard models of the form V_κ. We can now interpret the language multiple times, over V (the universal 'domain' for set theory) and over all the V_κ's. This yields collections of sentences that are true according to a model with 'domain' V, and according to all the models with domain V_κ. In the next step, I define the inherently absolutistic sentences with respect to V, and the relativistic sentences with respect to the V_κ. This way, we get the distinction from the semantics and not from outside.[50] The resulting picture has many advantages. To state the obvious first, bicontextualism gives us a more fine-grained position in the

48 Within this work, 'domains' (with quotation marks) indicates an ontologically neutral, non-standard reading of domains, according to which 'domains' are not entities, whereas domains (without quotation marks) simply refers to standard set-theoretic, entity-based domains.

49 The fact that I only consider the case where V contains all sets does not undermine the absolutistic aspect of the semantics. This is just for heuristic purposes.

50 In the end, the definition will be more complex, and I will discuss it in chapter 4 and chapter 5.

absolutism-relativism debate. So far, the positions between absolutists and relativists are typically developed in a very rigid and intransigent version.[51] According to this intransigent versions of absolutism and relativism, both deny any appeal of the opposed view. This, however, is too illiberal and leads to negative consequences at both extremes. Bicontextualism is intermediate between absolutism and relativism and avoids the disadvantages.

Second, splitting up the truths of the underlying language into Abs and Rel allows us to develop a more fine-grained approach as to how fundamental the respective claims are. In the following sections, we will see two applications of bicontextualism as a semantics for the language of truth and for the language of sets. The insights that we can get from these two applications are different. Roughly speaking, the difference is whether bicontextualism yields an implicit or an explicit approach to the respective concept. If the concept is truth, the insights we gain from bicontexualism are rather implicit, and if the concept is set, the insights are rather explicit. Let me explain this:

In the case of truth, we work with a language with a truth predicate and define, just like Kripke did, an extension of the truth predicate. However, I will develop a more dynamic picture as in the contextualist case. The result will be a sequence of extensions of the truth predicate, one for each context. Once we have that, we have also the class Abs containing inherently absolutistic truths. However, unlike the set-theoretic case, the truths will not primarily be concerned with the concept of truth itself. The underlying language will be the language of arithmetic. As in Kripke's theory of truth, we will be concerned with grounded sentences, i.e., sentences that are ultimately *grounded in facts*.[52] So the insights we gain about the concept truth will only be implicit in the construction. It tells us which sentences are so fundamental that they cannot be excluded from any contextual extension, but not which sentences about truth are fundamentally true.

Things are different in the case of set theory, as the object language quantifiers range over sets. So here the sentences that come out as inherently absolutistic truths are certainly among the most fundamental claims of the language of set theory. Sentences in that category have no countermodels, neither in the sense of absolutism nor in the sense of relativism. As such, inherently absolutistic sentences apply to any set whatsoever, and hence they convey some of the most fundamental features of the set-concept. In other words, in the set-theoretic case, we can use bicontextualism to extract core-properties of the set concept.

Related to that, bicontextualism respects the intended interpretation of the prima facie examples such as $\forall x\ x = x$. Similarly, in the context of set theory, it respects the fundamental nature of, say, the axiom of extensionality.

51 See in particular Glanzberg [49] and Lavine [81] for intransigent versions of relativism, and McGee [101] and Williamson [169] for intransigent versions of absolutism.

52 I assume that when we start in context 0 with a minimal fixed-point construction, than all sentences in context 0 are *directly grounded*. When we model the extension of the truth predicate for context 1, we do so by constructing the non-minimal fixed-point that results from taking the minimal fixed-point (from context 0), as the starting point. I assume that the sentences in context 1 and all subsequent contexts are *indirectly grounded*: they do not directly bottom out in facts about numbers, but can bottom out in facts about previous truth predicates. However, the succession of fixed-points will eventually terminate and so eventually, all indirectly grounded sentences bottom out in grounded sentences.

There is a special emphasis of the absolutist aspect of the semantics, which is philosophically motivated by the above considerations.

Finally, bicontextualism allows us to argue, contra Dummett, that paradoxes do not force us into a strict form of relativism. So, bicontextualism rejects the strong conclusion from the relativistic argument from paradox as too hasty, and restricts the scope of the conclusion to only the expansion sentences.

3 TECHNICAL BACKGROUND

In this section, I will provide the formal background for the following chapters. I will introduce the relevant notions of first-, second-, and higher-order logic, highlight the relevant metatheoretical differences, and give a rough sketch of plural logic, which will be important for the interpretation of higher-order expressions (section 3.1). I will introduce the technical background for type-free truth theories relevant to chapter 4 (section 3.2) and for set theory relevant to chapter 5 (section 3.3). Finally, I will present the RU approach as an absolutist-friendly alternative to model theory (section 3.4).

With the sole exception of the final part on the RU approach, none of the sections below is intended as a comprehensive presentation of the material. Rather, the aim is to gather the technical notions and theorems essential for the formal development in Chapters 4 and 5. Moreover, this section serves to introduce notational conventions.

3.1 FIRST-ORDER, HIGHER-ORDER, AND PLURAL LOGIC

3.1.1 *First-Order Logic*

Let $\mathcal{L}^1$ be a first-order language with identity $=$, negation $\neg$, conjunction $\wedge$, disjunction $\vee$, (material) conditional $\rightarrow$, as well as the existential quantifier $\exists$ and the universal quantifier $\forall$.[1] Unless stated otherwise, $\mathcal{L}^1$ contains countably many individual variables, denoted by x^0, y^0, z^0, possibly with indices x^0_1, x^0_2, x^0_3, countably many individual constants c^0, s^0, t^0, possibly with indices t^0_1, t^0_2, t^0_3, and countably many relation constants p^1, q^1, r^1, possibly with indices p^1_1, p^1_2, p^1_3. For simplicity, I assume that the language contains no function symbols. This is unproblematic, since functions can always be represented as relations that are univocal to the right. By Con^1 and Var^1, I denote the sets of $\mathcal{L}^1$-constants and $\mathcal{L}^1$-variables, respectively. The set of $\mathcal{L}^1$-terms, Ter^1, contains all $\mathcal{L}^1$-variables and $\mathcal{L}^1$-constants.

Theoretically, it suffices to take only a subset of the connectives and treat the others as defined abbreviations. For instance, with the basic set of connectives $\{=, \neg, \vee, \forall\}$, we can define $\varphi \rightarrow \psi$ as an abbreviation for $\neg(\varphi \vee \neg\psi)$, and $\exists x^0 \varphi(x^0)$ as an abbreviation for $\neg\forall x^0 \neg\varphi(x^0)$. The advantage of this more minimalist approach is that it simplifies the proofs of metatheoretical results as well as some definitions, e.g., the definition of satisfaction.

Atomic $\mathcal{L}^1$-formulae are of the form $x^0 = y^0$, where x^0 and y^0 are $\mathcal{L}^1$-terms, or of the form $r^1(t^0_1, \ldots, t^0_n)$, where r^1 is an n-ary $\mathcal{L}^1$-relation constant and $t^0_1, \ldots, t^0_n$ are $\mathcal{L}^1$-terms. The inductive clauses for well-formed formulae, as well as the notions of free and bound variables, are defined in the standard way. A

1 For a comprehensive introduction, see Boolos, Burgess, and Jeffrey [17] or Ebbinghaus, Flum, and Thomas [32].

sentence is a formula with no free variables. By For^1 and Sent^1, I denote the sets of $\mathcal{L}^1$-formulae and $\mathcal{L}^1$-sentences, respectively.

Note that I adopt the notational convention from the simple theory of types, according to which expressions denoting individuals are assigned type 0, and expressions denoting relations between individuals are assigned type 1. We will encounter expressions of higher types in the sections on higher-order logic below. Note further that I use the expressions 'type theory', 'simple theory of types', and 'higher-order logic' interchangeably. In particular, this means that the term 'type' and 'order' will be treated as synonyms.

A model for $\mathcal{L}^1$ is defined to be an ordered pair $\mathcal{M} = \langle M, I \rangle$, where M is a non-empty set, called the *domain*, and I is an interpretation function that assigns elements of M to $\mathcal{L}^1$-constants, and subsets of the n-fold Cartesian product M^n to n-ary relation constants. Whenever $\mathcal{M}$ is a model, I denote its domain by M. Satisfaction is defined relative to a model $\mathcal{M}$ and a variable assignment σ, which assigns to each $\mathcal{L}^1$-variable an element of M. The range of the first-order quantifiers is the domain M. An $\mathcal{L}^1$-sentence $\exists x^0 \varphi(x^0)$, with all free variables displayed, is said to be true relative to a model $\mathcal{M}$ and an assignment σ, written $\mathcal{M}, \sigma \models \exists x^0 \varphi(x^0)$, if and only if there exists an element in M that satisfies φ.

First-order logic has a number of metatheoretical results that distinguish it from second- and higher-order logic. Since these results will be important later on, I will briefly recall them here without proofs.

The most important result is certainly the adequacy theorem, which connects the syntactic notion of provability $\vdash$ with the semantic notion of logical consequence $\models$:

Theorem 3.1 (Adequacy)**.** For all $\Gamma \subseteq \mathsf{Sent}^1$ and $\varphi \in \mathsf{Sent}^1$,

$$\Gamma \models \varphi \text{ iff } \Gamma \vdash \varphi.$$

The adequacy theorem combines the soundness and completeness of first-order logic. The soundness theorem (right-to-left direction) tells us that if a sentence φ is provable from a set of assumptions Γ, then every model that satisfies all sentences in Γ also satisfies φ. The completeness theorem (left-to-right direction), on the other hand, tells us that if every model of Γ is also a model of φ, then there is a proof of φ from Γ.[2]

A direct corollary of the left-to-right direction of the adequacy theorem (completeness) is the compactness theorem:

Theorem 3.2 (Compactness)**.** Let $\Gamma \subseteq \mathsf{Sent}^1$. Then Γ is satisfiable if and only if every finite subset $\Gamma_{fin} \subseteq \Gamma$ is satisfiable.

A sentence φ is satisfiable if and only if there exists an $\mathcal{L}^1$-model $\mathcal{M}$ such that $\mathcal{M} \models \varphi$. Compactness tells us that a set of sentences Γ has a model if and only if every finite subset Γ_{fin} of Γ has a model. Thus, in order to determine whether an infinite set Γ has a model, it suffices to check whether each of its

2 Adequacy results are relative to a formal definition of provability, which always involves the specification of a proof system. Various options are available, such as Hilbert-style calculi, Gentzen-style systems, and others.

finite subsets does. The compactness theorem is a crucial tool for proving the existence of non-standard models of first-order theories.[3]

Finally, let me mention the Löwenheim-Skolem Theorem (LSK):[4]

Theorem 3.3 (Löwenheim-Skolem)**.**

1. **Downward:** Let $\mathcal{L}^1$ be a first-order language with $|\mathcal{L}^1| \geq \aleph_0$ (i.e., countable or larger), and let $\mathcal{M}$ be an $\mathcal{L}^1$-structure with $|M| \geq \aleph_0$, and $A \subseteq M$. Then, for every cardinality κ such that $|\mathcal{L}^1| + |A| \leq \kappa \leq |M|$, there exists an elementary substructure $\mathcal{M}_0 \lhd \mathcal{M}$ such that $A \subseteq M_0$ and $|M_0| = \kappa$.
2. **Upward:** Let $\mathcal{L}^1$ be a first-order language with $|\mathcal{L}^1| \geq \aleph_0$, and let $\mathcal{M}$ be an $\mathcal{L}^1$-structure with $|M| \geq \aleph_0$. Then, for every cardinality κ such that $\kappa \geq \max(|M|, |\mathcal{L}^1|)$, there exists an $\mathcal{L}^1$-structure $\mathcal{M}^*$ such that $\mathcal{M} \lhd \mathcal{M}^*$ and $|M^*| = \kappa$.

Thus, the LSK theorem shows that both smaller and larger models can be found for any infinite model of a first-order theory. The downward version tells us that whenever we have a model of arbitrary infinite cardinality, we can shrink it to a model of smaller size, in particular to one of cardinality $|\mathcal{L}^1|$ when $A = \emptyset$. The idea is that it suffices to interpret the (countably many) non-logical symbols—individual constants and relation constants—to determine the size of the domain. The upward version tells us that we can always find models of arbitrarily larger cardinality, since we can freely expand the domain while suitably extending the interpretation function.

The LSK and the compactness theorem are commonly regarded as *limitative* results of first-order logic. In light of these results, it becomes evident that first-order logic cannot uniquely determine the size of its models—in particular, it cannot distinguish between different notions of infinity.[5] These limitations have far-reaching consequences for the semantics of first-order logic. Consider, for instance, first-order real analysis: by Cantor's theorem, we know that the cardinality of the continuum exceeds $\aleph_0$. Nevertheless, by the downward version of LSK, first-order real analysis admits a model of countable size. The situation is even more striking in the case of set theory: starting with any transitive model of ZFC, the downward LSK guarantees the existence of a *countable transitive* model of ZFC. These results are known as *Skolem's paradox*.[6]

3.1.2 *Higher-Order Logic*

Second-order logic extends first-order logic by introducing quantification not only over individual variables but also over predicates, relations, and functions (although, as in the first-order case, I will omit functions here for simplicity).[7]

3 I will sketch a compactness argument for non-standard models of set theory in section 3.3.3.

4 Note that I use $\lhd$ for elementary substructurehood because the usual $\preceq$ is already in use (section 3.1.3).

5 This will be important for the motivation to incorporate higher-order object languages in section 4.1.

6 Skolem's paradox is, strictly speaking, not a paradox, as no contradiction arises from the existence of countable transitive models of ZFC. Nonetheless, these results are often considered counterintuitive and philosophically troubling. For discussion, see Bays [5].

7 For a comprehensive introduction to second-order logic, see Bell [9] or Bacon [2].

Let $\mathcal{L}^1$ be a first-order language. The language of second-order logic, denoted $\mathcal{L}^2$, is a proper extension of $\mathcal{L}^1$, that is, $\mathcal{L}^2 \supset \mathcal{L}^1$. In other words, $\mathcal{L}^2$ contains all the logical connectives, individual constants, and variables of $\mathcal{L}^1$—the so-called type-0 terms—as well as all the relation constants (type-1 constants) of $\mathcal{L}^1$.

Crucially, $\mathcal{L}^2$ also introduces countably many relation variables (type-1 variables), typically denoted x^1, y^1, z^1, possibly indexed as x^1_1, x^1_2, x^1_3, and so on. The key feature of second-order logic is that quantifiers can bind these type-1 variables, i.e., second-order logic allows quantification over relations. Throughout, I will mainly follow a type-theoretic naming convention, referring to individual constants and variables as type-0 terms, relation constants and variables as type-1 terms, and so on. Occasionally, however, I will use the more familiar expressions 'second-order variable' for type-1 variables, 'third-order variable' for type-2 variables, and so on. The reader should keep in mind that these are equivalent ways of referring to the same objects.

The atomic formulae of $\mathcal{L}^2$ consist of all atomic $\mathcal{L}^1$-formulae together with expressions of the form $x^1(t^0_1, \ldots, t^0_n)$, where x^1 is an n-ary type-1 variable and each t^0_i is a type-0 term. $\mathcal{L}^2$-quantifiers can not only bind type-0 variables but also type-1 variables, so $\mathcal{L}^2$-quantification is $\mathcal{L}^1$-quantification plus quantification over type-1 variables. The set of well-formed formulae of $\mathcal{L}^2$ is defined recursively, following the usual clauses for $\mathcal{L}^1$-formulae but augmented by an additional clause to account for second-order quantification:

> If φ is an $\mathcal{L}^2$-formula and x^1 is a type-1 variable, then $\forall x^1, \varphi$ is also an $\mathcal{L}^2$-formula.

I will write Con^2 and Var^2 for the sets of constants and variables of $\mathcal{L}^2$, respectively. The set of $\mathcal{L}^2$-terms, denoted Ter^2, contains all constants and variables of $\mathcal{L}^2$, that is, all type-0 and type-1 terms.

More generally, $(n{+}1)^{\text{th}}$-order logic is an extension of n^{th}-order logic that allows quantification over type-n terms. If $\mathcal{L}^n$ is an n^{th}-order language, the language of $(n{+}1)^{\text{th}}$-order logic, $\mathcal{L}^{n+1}$, is a proper superlanguage, $\mathcal{L}^{n+1} \supset \mathcal{L}^n$, i.e., $\mathcal{L}^{n+1}$ contains all the connectives and terms of $\mathcal{L}^n$. In addition, $\mathcal{L}^{n+1}$ contains countably many constant and variable type-n terms.[8]

In the case of n^{th}-order logic for $n > 2$, we face a choice between *strict* and *cumulative* type theory. This distinction concerns whether *cross-type predication* is allowed. According to strict type theory, a predication of the form $c^n(x^i)$ is well-formed if and only if $n = i + 1$. Thus, predication is allowed only in cases such as $c^n(x^{n-1})$, $c^4(x^3)$, and so on. Accordingly, the atomic $\mathcal{L}^{n+1}$-formulae consist of all atomic $\mathcal{L}^n$-formulae, along with all formulae of the form $x^n(t^{n-1}_1, \ldots, t^{n-1}_m)$, where x^n is an m-ary (variable or constant) type-n term, and $t^{n-1}_1, \ldots, t^{n-1}_m$ are (variable or constant) type-$(n{-}1)$ terms.

In contrast, cumulative type theory is more liberal: Predication of the form $c^n(x^i)$ is well-formed if and only if $i < n$, that is, whenever the predicated term

8 Note the difference in the transition from first- to second-order logic and from n^{th}- to $(n{+}1)^{\text{th}}$-order logic: $\mathcal{L}^1$ contains constant and variable type-0 terms *and* constant type-1 terms. When moving to $\mathcal{L}^2$, we add variable type-1 terms and allow quantifiers to bind them. By contrast, $\mathcal{L}^n$ contains only constant and variable terms of type $n{-}1$, but no constant type-n terms. Thus, when extending to $\mathcal{L}^{n+1}$, we introduce both variable *and* constant type-n terms, and permit quantifiers to bind them.

is of lower type than the predicate itself. Thus, under cumulative type theory, expressions like $c^8(x^2)$ and $c^n(x^0)$ (for $n > 0$) are legitimate predications. Consequently, the atomic $\mathcal{L}^{n+1}$-formulae include all atomic $\mathcal{L}^n$-formulae, plus all formulae of the form $x^n(t^i_1, \ldots, t^j_m)$, where x^n is an m-ary (variable or constant) type-n term, and each $t^i_1, \ldots, t^j_m$ is a (variable or constant) term of type i or j, for some $i, j < n$.

In either case, the definition of complex $\mathcal{L}^{n+1}$-formulae proceeds by extending the recursive clauses for $\mathcal{L}^n$-formulae with the following additional clause:

> If φ is an $\mathcal{L}^{n+1}$-formula and x^n is a variable type-n term, then $\forall x^n \varphi$ is an $\mathcal{L}^{n+1}$-formula.

Second-order logic has two distinct semantics. Since relations are interpreted as sequences of elements from the domain, a natural interpretation of second-order variables is that they are assigned to elements of the powerset of the domain, $\mathcal{P}(M)$. In this interpretation, second-order quantifiers range over $\mathcal{P}(M)$. Specifically, a $\mathcal{L}^2$-sentence of the form $\exists x^1 \varphi(x^1)$, with all free variables displayed, is true relative to a model $\mathcal{M}$ and an assignment σ if and only if there exists some element in $\mathcal{P}(M)$ that satisfies φ. This approach is known as *full semantics*, as it considers the *full* powerset as the quantifier range.

Alternatively, one may restrict the range of the second-order quantifiers to a specific subset $R \subseteq \mathcal{P}(M)$. In this case, a $\mathcal{L}^2$-sentence of the form $\exists x^1 \varphi(x^1)$ is true relative to a model $\mathcal{M}$ and an assignment σ if and only if there is some element in $R \subseteq \mathcal{P}(M)$ that satisfies φ. This restricted interpretation is referred to as *Henkin semantics*.

In third- and higher-order logic, we encounter a similar distinction. In full semantics, third-order variables are interpreted over the powerset of the powerset of the domain, whereas in Henkin semantics, they are interpreted over a subset of the powerset of the powerset of the domain. This pattern continues in fourth- and higher-order logic, generalizing in an analogous way.

A model of full second-order logic is of the form $\mathcal{M}^2 = \langle M, \mathcal{P}(M), I \rangle$, where $\mathcal{P}(M)$ is the powerset of the domain M, and I is an interpretation function. A model of full third-order logic has the form $\mathcal{M}^3 = \langle M, \mathcal{P}(M), \mathcal{P}(\mathcal{P}(M)), I \rangle$, where $\mathcal{P}(\mathcal{P}(M))$ is the powerset of the powerset of M. A Henkin-model of second-order logic is of the form $\mathcal{M}^2_H = \langle M, R, I \rangle$, where $R \subseteq \mathcal{P}(M)$. Similarly, a Henkin-model of third-order logic is of the form $\mathcal{M}^3_H = \langle M, R, R', I \rangle$, where $R \subseteq \mathcal{P}(M)$ and $R' \subseteq \mathcal{P}(\mathcal{P}(M))$. As expected, this pattern generalizes in a straightforward manner for higher-order logic.

The choice of semantics has far-reaching consequences and marks the key metalogical difference between first-order and higher-order logic. Specifically, any n^{th}-order logic for $n > 1$ is complete, compact, and has the Löwenheim-Skolem property when interpreted through a Henkin-style semantics. Under this approach, higher-order logic can always be reduced to multisorted first-order logic.[9] However, when interpreted through full semantics, higher-order logic loses all these metalogical properties. Second- and higher-order logic under full semantics is neither complete nor compact, and does not have the Löwenheim-Skolem property.

9 See Shapiro [149, p. 137 ff.]

3.1.3 *(The) Plural (Interpretation of Higher-Order) Logic*

Plural logic is an extension of first-order logic that introduces plural terms (such as xx for plural variables and cc for plural constants) and plural quantifiers, including $\forall xx$ (*for any things* xx) and $\exists xx$ (*there are some things* xx).[10]

From a syntactic perspective, plural logic extends first-order logic by adding plural constants and variables, along with the binary predicate $\preceq$ (*is one of*). Consequently, the terms in plural logic consist of the terms from first-order logic, augmented by plural terms. The definition of well-formed formulae in plural logic includes all the clauses from first-order logic, supplemented by the following:

- $t \preceq tt$, where t is a singular term and tt is a plural term;
- $\forall xx\varphi$, where xx is a plural term and φ is a formula.

Formally, plural quantification corresponds to monadic second-order quantification, but with distinct notation. Instead of using $x^1(x^0)$, plural logic uses $t \preceq tt$. Thus, the semantics for plural quantification are analogous to the semantics of second-order quantification.

A key distinction between plural logic and second-order logic lies in the denotation of plural expressions, such as xx or cc. Unlike second-order expressions, which, according to standard semantics, denote set-theoretic entities that represent collections of objects, plural expressions simultaneously denote a multitude of objects from the first-order domain. Thus, plural expressions do not introduce a new ontological category (like sets) but rather a new kind of reference, known as *plural denotation*.[11]

To illustrate, consider the predicate 'is prime'. In set-theoretic semantics, its denotation is a set such as $P = 2, 3, 5, 7, ...$ Under the plural interpretation, by contrast, 'is prime' refers directly to the primes themselves—without invoking a set. This feature exemplifies what has traditionally been called the *ontological innocence of plural reference*, which gois back to the Boolos early work on plural logic.[12]

Although a set-based semantics for plural logic can be provided,[13] this section will focus exclusively on a *plurality-based semantics*. Like second-order logic, plural logic can be interpreted with either *full* or *Henkin* semantics.

Let a *superplurality* be a plurality of pluralities, and a *subplurality* be a subcollection of a plurality, analogous to a subset in set theory. A full superplurality

10 The plural interpretation was developed by Boolos, see Boolos [16, 12]. For more recent treatments, see Florio and Linnebo [42, 41, 43], Oliver and Smiley [114], and Linnebo [86].

11 See Rayo [132].

12 See Boolos [16, 12]. Note that the ontological innocence of plural quantification remains a controversial topic. For instance, Resnik [139] argues that plural expressions denote set-like entities and are therefore not ontologically innocent. Linnebo [87], while not strictly denying the ontological innocence of plural quantification *tout court*, contends that true instances of *impredicative* plural comprehension bear as much ontological commitment as their set-theoretic counterparts. Florio and Linnebo [42] argue that a plural Henkin semantics reveals the ontological commitment of a theory T by highlighting the minimal sub-superplurality of the first-order domain that every Henkin model of T must contain. On the other hand, Boolos [16, 12] and Uzquiano [159] use plurals to develop an eliminativist view about classes.

13 See Florio and Linnebo [43, Ch. 7.2].

covers all subpluralities of a given plurality, similar to how the full powerset covers all subsets in set theory. Under full semantics, plural quantifiers range over the full superplurality of the first-order domain, while under Henkin semantics, they range over a subplurality of the full superplurality, mirroring the way second-order quantifiers range over a subset of the full powerset in Henkin semantics.

Notably, plural logic shares the same semantic metatheorems as second-order logic. When interpreted with Henkin semantics, it is complete, compact, and has the Löwenheim-Skolem property. In contrast, under full semantics, plural logic lacks all of these properties.

Plural quantification can be generalized to higher levels, analogous to the extension of second-order quantification to even higher orders.[14] This leads to increasingly complex constructs, such as superpluralities (xxx), which denote pluralities of pluralities, and superduperpluralities ($xxxx$), which represent pluralities of pluralities of pluralities, and so on. Each successive level introduces a new kind of reference, rather than a new ontological category. However, it is important to note that these higher-level plural constructs remain the subject of ongoing philosophical debate. The motivation for expanding the hierarchy to higher-level plurals is not as clear-cut as the motivation for introducing plural expressions into first-order logic.[15]

Moreover, the determinacy of reference for higher-level plural expressions seems to threaten the ontological innocence even of plural expressions themselves.[16] However, there has been progress in defending higher-level plurals. In particular, Grimau [55] has argued that objections to higher-level plurals can be *mutatis mutandis* transformed into objections to plural expressions in the first place. If these considerations are correct, it is not clear why one would adopt plural expressions but not higher-level plurals. Consequently, one would have to adopt either both or neither.[17]

14 See in particular Rayo [132] and Oliver and Smiley [114].

15 The motivation to add plural resources to first-order logic stems from considerations in natural language. Expressions like ´the students', 'Luca's friends', and 'Kim, Robin, and Mika' are natural language plural expressions. However, finding natural language examples for higher-level plurals is more difficult. For instance, Linnebo and Nicolas [92] argue that expressions like 'Russell and Whitehead, and Hilbert and Bernays' (p. 189), which seem to be superplurals, can actually be reduced to plural expressions. The examples they provide for English superplurals include 'These people and those people' and 'The square things and the blue things' (p. 193). One should note that even if these arguments are successful, Linnebo and Nicolas [92] have only motivated the transition from plural logic to superplural logic, which would correspond to the transition from second-order logic to third-order logic. The case for motivating an entire hierarchy of higher-level plural logic has yet to be made.

16 If an expression denotes a plurality of objects (say, *dogs* denotes some dogs), then what, exactly, is the denotation of a higher-level plural? For instance, assume that *dogses* is an English higher-level plural expression denoting a plurality of pluralities of dogs. This seems to commit us to a reading of the plural expression that is not ontologically innocent. Scholars skeptical of the ontological innocence of plural quantification might be even more skeptical of the ontological innocence of higher-level plural quantification (see footnote 12).

17 For instance, arguments against higher-level plurals based on paraphrasing them as plural expressions are not always available, and objections based on reification raise the question of why reification should be used for superplurals but not already for plurals. Thus, if one accepts these objections to higher-level plurals, one must also explain why these arguments are not equally conclusive against plurals themselves.

In this work, I follow the tradition of plural logic developed by Boolos, rather than adopting the framework of critical plural logic (CPL) introduced by Florio and Linnebo [42, 41, 43]. This choice is motivated by the broader philosophical aims of the project, particularly the aim of supporting an absolutist view of quantification.

The debate between Boolos's plural logic and CPL concerns a tension between three principles: Universal Singularization (the idea that every plurality forms an object), Unrestricted Plural Comprehension (the idea that any condition determines a plurality), and Absolute Generality. These three cannot all be maintained simultaneously. From Unrestricted Plural Comprehension, we get a universal plurality; from Universal Singularization, this plurality forms a single object. But then Absolute Generality is lost, as the object representing the universal plurality cannot itself be among the things it ranges over.

Given this conflict, one must choose which principle to reject. Florio and Linnebo choose to preserve Universal Singularization at the expense of Unrestricted Plural Comprehension. The resulting logic—CPL—permits the singularization of pluralities into sets, but restricts which conditions count as defining pluralities. Florio and Linnebo show that critical plural logic, when supplemented with suitable assumptions, allows for the reconstruction of principles that correspond closely to the axioms of ZFC.

My approach takes the opposite route. I reject Universal Singularization and maintain Unrestricted Plural Comprehension. In my view, it is entirely natural to suppose that any condition defines a plurality—a point Florio and Linnebo themselves acknowledge as a viable option.[18] However, the view adopted here denies that every such plurality forms a set. Following Hewitt [66], I hold that while some pluralities form sets, others do not. As Hewitt puts it, this position casts doubts on "the availability of any informative principle which will settle which pluralities form sets."[19]

On this view, plural logic is not a theory about which sets exist, but a method for interpreting higher-order discourse without ontological reification. The existence of sets is a substantive question addressed by set theory, not something that plural logic alone determines. Accordingly, Boolos' plural logic—permissive in its comprehension and free of singularizing commitments—offers a more appropriate framework for my purposes than CPL.

Within this work, I will not explicitly use plural logic as a formal system. However, it is a key component—indeed, the central interpretative assumption—for the absolutist-friendly understanding of higher-order expressions that underlies the philosophical framework developed here. While I continue to use the notational conventions of type theory, this usage should be understood as a heuristic device. The intended interpretation of higher-order variables and quantifiers is thoroughly plural: second-order variables are interpreted as ranging over pluralities of objects, third-order variables as ranging over pluralities of pluralities, and so on.

This plural interpretation plays a foundational role for unrestricted quantification in a higher-order setting. In particular, it provides a way to talk about

18 See Florio and Linnebo [43, p. 276].

19 Hewitt [66, p. 311].

"absolutely everything" without reifying the totality of all sets or properties as a single object or domain. A more detailed discussion of an absolutist-friendly plural interpretation of higher-order expression is given in section 6.1.

It is worth noting, however, that although I follow Boolos's interpretation of plural logic, my approach goes beyond what has been suggested by Boolos's, as he himself was skeptical about the legitimacy of higher-order plurals.[20] So, the plural interpretation used here builds on the ontological innocence of plural expressions insights as a means to interpret higher-order languages in an absolutist-friendly way. While Boolos may have doubted the viability of higher-order pluralities, I contend–drawing on defenses such as Grimau [55] mentioned in footnote 16 and Rayo [132] and Rayo and Uzquiano [135]—that this view can be coherently developed.

3.2 THEORIES OF TYPE-FREE TRUTH

This section outlines the relevant background for the discussion of truth-theoretical bicontextualism. In particular, I provide the background for coding (section 3.2.1), inductive definitions (section 3.2.2), Kripke's truth theory (section 3.2.3), contextualism (section 3.2.4), and finally, a brief note on the axiomatization of Kripke's truth theory (section 3.2.5).

3.2.1 *Coding and Acceptable Structures*

I already touched on the topic of coding in section 2.2.2. Let me make it precise now. Coding is required to define self-referential sentences such as the liar λ, which says of itself that it is not true: $\lambda :\leftrightarrow \neg \mathsf{Tr}(\lfloor \lambda \rfloor)$.

Truth-theoretic research usually takes the standard language of first-order arithmetic $\mathcal{L}^1_{\mathbb{N}} = \{0, S, +, \times, <\}$ as a toy model of natural language. This language contains 0 as the constant symbol for zero, S, $+$, and $\times$ as function symbols for the successor, addition, and multiplication functions, respectively, and a binary predicate $<$ for the natural ordering on $\mathbb{N}$.

I expand $\mathcal{L}^1_{\mathbb{N}}$ by adding a truth predicate, symbols for primitive recursive functions, and second-order parameters:

Definition 3.4. The language $\mathcal{L}_{\mathsf{Tr}}$ is defined as an expansion of the first-order language $\mathcal{L}^1_{\mathbb{N}}$ by:

1. Function symbols for all primitive recursive functions needed to prove the strong diagonalization lemma;
2. A fresh unary predicate symbol Tr to serve as a truth predicate;
3. *Free second-order parameters* in the language.[21]

The behavior of these symbols is governed by the axiomatic system of Dedekind-Peano Arithmetic (PA) together with suitable axioms for the primi-

20 See Boolos [16, 12].

21 This is needed for the theory of inductive definitions (section 3.2.2). Equivalently, one may assume that for *any* relation $B \subseteq \mathbb{N}^m$, we can expand the signature with a relation constant.

tive recursive functions. The resulting system can express its own syntax via a sufficiently strong coding schema, defined as an injective mapping

$$\lfloor \cdot \rfloor : \mathcal{L}_{\mathsf{Tr}} \to \mathbb{N},$$

which assigns to every expression its unique code. Consequently, for any $\mathcal{L}_{\mathsf{Tr}}$-sentence φ, $\lfloor \varphi \rfloor$ is the name of φ.[22] Moreover, it can prove (a strong form of) the diagonal lemma, which implies the existence of closed terms t_φ for each formula φ such that $t_\varphi = \lfloor \varphi(t_\varphi) \rfloor$ is provable from the base theory.[23]

To deal with free second-order parameters, the satisfaction relation $\models$ is defined as in full (or Henkin) second-order logic. That is, if a formula φ of the expanded language contains free relation variables, its truth in a structure $\mathcal{M}$ is evaluated relative to a second-order assignment σ which maps each free relation variable to the appropriate subset or relation over the domain of $\mathcal{M}$. However, the language does not allow to quantify over second-order variables.

From here, I can proceed in one of two ways. I can either proceed axiomatically, proving the strong diagonal lemma in the axiomatic system outlined above and defining a concrete coding scheme, or I can proceed model-theoretically, specifying the conditions a structure must satisfy in order to be able to define such a coding scheme, and then dealing only with structures that are *acceptable* in this sense. I take the model-theoretic route.

Acceptable structures can be found in Moschovakis [106, Ch. 5]. However, a slight modification is necessary: Moschovakis' definitions require that acceptable structures contain an isomorphic copy of the natural numbers. The concept of isomorphism between two structures demands that both structures are defined over the same signature. The arithmetical signature provided earlier contains function symbols $(S, +, \times)$. For simplicity, however, I focus on languages that include only relational symbols (see definitions 3.31, 4.1, and 5.1). Therefore, I adapt Moschovakis' original definitions as suggested in McGee [102, Ch. 1], where structures are considered acceptable if they can interpret the natural numbers structure. I begin with Moschovakis' original definition and then describe the necessary adjustments.

Definition 3.5. A *coding scheme* for a structure $\mathcal{M}$ is a triple $\langle N, S_N, \lfloor \cdot \rfloor \rangle$ where

- $N \subseteq M$, and $\langle N, S_N \rangle \cong \langle \mathbb{N}, S \rangle$, and
- $\lfloor \cdot \rfloor : \bigcup_{n \in \omega} M^n \to M$ is an injection from the set of all finite sequences of elements of M into M,

such that the following are definable:

- $\mathsf{seq}(x)$, the set of codes of finite sequences;

22 For details, see Boolos, Burgess, and Jeffrey [17], Dalen [26], and Hájek and Pudlák [56]. As mentioned earlier, I use $\lfloor \cdot \rfloor$ for coding because the usual $\ulcorner \cdot \urcorner$ is already in use (section 3.4).

23 For self-reference, I rely on the *strong* rather than the *weak* diagonalization lemma. The weak diagonal lemma implies only the material equivalence $\varphi \leftrightarrow \varphi(\lfloor \psi \rfloor)$ of a formula and its diagonalization. By the strong diagonal lemma, the terms denoting the diagonalization exist in the language, and the equality $t_\varphi = \lfloor \varphi(t_\varphi) \rfloor$ is provable. See Heck [64] and Picollo [121] for details.

- $\mathsf{lh}(x) = \begin{cases} 0, & \text{if } \neg\mathsf{seq}(x) \\ m, & \text{if } \mathsf{seq}(x) \wedge x = \lfloor x_1, ..., x_m \rfloor \end{cases}$
- $\mathsf{q}(x,i) = \begin{cases} x_i, & \text{if } x = \lfloor x_1, ..., x_m \rfloor \wedge 1 \leq i \leq m \\ 0, & \text{otherwise.} \end{cases}$

We call a structure $\mathcal{M}$ *acceptable* if it admits a coding scheme.

Coding is primarily used in the truth-theoretic part of this work (chapter 4), where the object languages are all countable, so it suffices to define a coding scheme as an injection into (an isomorphic copy of) $\mathbb{N}$.[24]

A relative interpretation of the natural numbers in a target model is, roughly speaking, a translation function from $\mathcal{L}_{\mathbb{N}}$ into the target language, such as the language of set theory $\mathcal{L}_{\in}$, that preserves theoremhood. More precisely, let $\mathcal{M}$ be an $\mathcal{L}_{\in}$-structure. First, we identify $\mathcal{L}_{\in}$-formulae that represent the relevant $\mathcal{L}_{\mathbb{N}}$-notions inside $\mathcal{M}$. For instance, we use the $\mathcal{L}_{\in}$-formulae $N(x)$ to represent the copy of the domain of $\mathbb{N}$ in $\mathcal{M}$, $Z(x)$ to represent 0, $S(x,y)$ to represent the successor function, $A(x,y,z)$ to represent the addition operation, $M(x,y,z)$ to represent the multiplication operation, and $L(x,y)$ to represent the less-than relation in $\mathcal{M}$.

The translation of $\mathcal{L}_{\mathbb{N}}$ to $\mathcal{L}_{\in}$ is defined only for *unnested formulae*, i.e., formulae in which the only terms are variables, individual constants, and functions that take only variables as arguments. To translate a formula, replace all occurrences of non-logical symbols accordingly—replace $x = 0$ with $Z(x)$, $a \times b = c$ with $M(a,b,c)$, and so on. Then, relativize each quantifier as follows: replace $\exists x \varphi(x)$ with $\exists x(N(x) \wedge \varphi(x))$. A *relative interpretation* is defined as the translation of all PA-axioms together with additional clauses that ensure that there is only one element satisfying $N(x) \wedge Z(x)$ and that the functions defined by $S(x,y)$, $A(x,y,z)$, and $M(x,y,z)$ are defined on $N(x)$. A $\mathcal{L}_{\in}$-theory T relatively interprets PA if and only if each of these sentences of the relative interpretation is a T-theorem.[25]

For definition 3.5, this means that a structure $\mathcal{M}$ is acceptable if it contains a substructure that can relatively interpret the natural numbers and can represent (relational versions of) seq, lh, and q. From now on, we will refer to structures as 'acceptable' in this sense. Moreover, I assume that the standard language for truth contains at least $\in$ and the truth predicate Tr.

3.2.2 *Inductive Definitions*

The formal background for Kripke's model-theoretic construction (see section 2.2.2) is the theory of inductive definitions and monotone operators.[26] An

24 Coding schemes for large uncountable languages can be defined as functions into large structures, such as $\langle H(\kappa), \in\restriction_{H(\kappa)}\rangle$, where κ is the cardinality of the language, and $H(\kappa)$ is the set of all sets hereditarily of cardinality less than κ, see Dickmann [28].

25 For more on relative interpretations, see McGee [102, ch. 1], Hodges [70, ch. 5], or Button and Walsh [18, ch. 5].

26 The classic reference for the theory of inductive definitions is Moschovakis [106]. Alternatively, see Aczel [1], or Pohlers [124, ch. 6].

inductive definition generates a set of objects as the smallest set containing some initially given objects (base case) by iterated application of some specific closure rules. For instance, we can define the set of natural numbers as the smallest set that contains 0 (base case) and is closed under the successor operation. In abstract terms, an inductive definition consists of a collection of clauses of the form

$$\forall a(a \in A \rightarrow a \in B) \Rightarrow b \in B. \tag{3.1}$$

Consider the definition of the set B of well-formed formulae (wffs) of propositional logic. Let $\mathcal{L} = \{\wedge, \vee, \neg, \rightarrow, \leftrightarrow, P_0, P_1, P_2, ...\}$ be the language of propositional logic. B is defined inductively via a base case:

- **Base case:** If P_n is a propositional atom, then P_n is a wff;

together with a $\neg$- and a $\circ$-closure rule (for $\circ \in \{\wedge, \vee, \rightarrow, \leftrightarrow\}$) in form of 3.1:

- **Closure rule $\neg$:** If P_n is a wff, then so is $\neg P_n$;
- **Closure rule $\circ$:** If P_n and P_m are wff, then so is $P_n \circ P_m$.

The base clause specifies some initial elements A that are in the set B of $\mathcal{L}$-wffs, in this case, the propositional atoms. This corresponds to the antecedent of eq. (3.1). In the inductive clauses, b is an expression of the form $\neg P_n$ or of the form $P_n \circ P_m$. The inductive clause for $\neg$ can be read as 'If the propositional atom P_n is a wff (i.e., $\in B$), then so is $\neg P_n$'. This corresponds to the consequent of eq. (3.1). We can simplify 3.1 and write it as $A \Rightarrow b$, stating that if some initial elements are taken as the starting point of an inductive definition (i.e., are in A), then b belongs to the inductively generated set. This motivates the following definition:

Definition 3.6. An inductive definition is a collection of clauses $\mathcal{A}$ of the form $A \Rightarrow b$. A set B satisfies an inductive definition $\mathcal{A}$ (alternatively: B is closed under $\mathcal{A}$) if

$$((A \Rightarrow b) \in \mathcal{A} \wedge A \subseteq B) \Rightarrow b \in B.$$

A monotone operator is a function on the powerset of a given set that satisfies the monotonicity requirement:

Definition 3.7. Let B be an infinite set. An operator is a function on the powerset of B,

$$\Gamma : \mathcal{P}(B) \rightarrow \mathcal{P}(B),$$

Γ is *monotone* if it preserves set-inclusion, i.e., for $B_0, B_1 \in \mathcal{P}(B)$,

$$B_0 \subseteq B_1 \Rightarrow \Gamma(B_0) \subseteq \Gamma(B_1).$$

Inductive definitions and monotone operators can be viewed as two sides of the same coin. Consider the application of a monotone operator to a given set B as a method for closing under the clauses contained in the corresponding

inductive definition. Specifically, for an inductive definition $\mathcal{A}$ and a set $B_0 \in \mathcal{P}(B)$, a monotone operator on B can be interpreted as

$$\Gamma_{\mathcal{A}}(B_0) = \{b \in B \mid \exists(A \Rightarrow b) \in \mathcal{A} \wedge (A \subseteq B_0)\}.$$

In other words, the image of B_0 under the monotone operator $\Gamma_{\mathcal{A}} : \mathcal{P}(B) \to \mathcal{P}(B)$ consists of all $b \in B$ such that there exists a condition $(A \Rightarrow b)$ in $\mathcal{A}$ with $A \subseteq B_0$. Moreover, this clearly satisfies the monotonicity criterion: for $B_0 \subseteq B_1$, any element added by clauses involving elements in B_0 is also added by clauses involving elements in B_1, which ensures that the operator is monotone.

We can think of a monotone operator Γ on B (or, equivalently, an inductive definition) as generating an ordinal-indexed, possibly transfinite, sequence of increasing subsets of B, referred to as the *stages of the inductive definition*, denoted by I^{α}_{Γ}. This sequence eventually leads to the *set inductively defined by* Γ, denoted by I_{Γ}. The process can be formalized using transfinite recursion, so that for an ordinal α,

$$I^{\alpha}_{\Gamma} := \Gamma\Big(\bigcup_{\beta<\alpha} I^{\beta}_{\Gamma}\Big) \quad \text{and} \quad I^{<\alpha}_{\Gamma} = \bigcup_{\beta<\alpha} I^{\beta}_{\Gamma},$$

such that $I^{\alpha}_{\Gamma} = \Gamma(I^{<\alpha}_{\Gamma})$ for every α. Additionally, we define

$$I_{\Gamma} := \bigcup_{\alpha \in Ord} I^{\alpha}_{\Gamma}.$$

The resulting sequence is then given by

$$\begin{aligned} I^0_{\Gamma} = \Gamma(\varnothing) \subseteq I^1_{\Gamma} = \Gamma(\Gamma(\varnothing)) \subseteq I^2_{\Gamma} &= \Gamma\left(I^0_{\Gamma} \cup I^1_{\Gamma}\right) \\ &= \Gamma\left(\Gamma(\varnothing) \cup \Gamma(\Gamma(\varnothing))\right) \subseteq I^3_{\Gamma} \dots \end{aligned}$$

For cardinality reasons, this process reaches a *fixed-point* where $I^{\alpha}_{\Gamma} = I^{<\alpha}_{\Gamma}$ (and consequently, $\Gamma(B') = B'$ for some $B' \in \mathcal{P}(B)$), with α being bounded from above by the cardinality of the set on which Γ operates. In other words, an operator on B generates a sequence that reaches a fixed-point in at most $|B|$-many steps:[27]

Theorem 3.8. Let Γ be a monotone operator on B. Then, for some ordinal $\alpha \leq |B|$, we have $I_{\Gamma} = I^{\alpha}_{\Gamma} = I^{<\alpha}_{\Gamma}$. Moreover,

$$I_{\Gamma} = \bigcap\{B_0 \in \mathcal{P}(B) \mid \Gamma(B_0) = B_0\},$$

which is the smallest fixed-point.

Proof. I follow Moschovakis [106, Th. 1A.1]. Assume for the sake of contradiction that $I^{<\alpha}_{\Gamma} \subset I^{\alpha}_{\Gamma}$ for every $\alpha < |B|^{+}$. This implies that at each step of the inductive definition, new elements are added, so that for $|B|^{+}$-many steps, new elements are introduced at each stage. For each step, select one of the new elements, i.e., $x_{\alpha} \in I^{\alpha}_{\Gamma} \setminus I^{<\alpha}_{\Gamma}$, and define the set

$$Y = \{x_{\alpha} \mid \alpha < |B|^{+}\}.$$

27 See Moschovakis [106, Th. 1A.1] or Pohlers [124, Obs. 6.2.4 and Lemma 6.3.2].

This would yield a subset of B with cardinality $|B|^+$, which would absurdly exceed the cardinality of B. Hence, there exists some $\alpha < |B|^+$ such that $I_\Gamma^\alpha = I_\Gamma^{<\alpha}$. For any $\gamma > \alpha$, we then have $I_\Gamma^\gamma = I_\Gamma^\alpha$, and thus $I_\Gamma^\alpha = I_\Gamma$. Therefore, I_Γ is a fixed-point, as $\Gamma(I_\Gamma) = \Gamma(I_\Gamma^\alpha) = \Gamma(I_\Gamma^{<\alpha}) = I_\Gamma^\alpha = I_\Gamma$.

Now, consider any fixed-point B' of Γ, i.e., any $B' \in \mathcal{P}(B)$ such that $\Gamma(B') = B'$. We will show that $I_\Gamma^\xi \subseteq B'$ for all ordinals ξ by transfinite induction on ξ. This will imply that I_Γ is the smallest fixed-point. The induction hypothesis is that $I_\Gamma^{<\xi} \subseteq \Gamma(B')$. Then,

$$I_\Gamma^\xi = \Gamma(I_\Gamma^{<\xi}) \subseteq \Gamma(\Gamma(B')) = B',$$

by monotonicity. Thus, if $x \in I_\Gamma$, then there exists an ordinal α such that $x \in I_\Gamma^\alpha$. By the above reasoning, this implies $x \in B'$. Since B' was arbitrary, we conclude that $x \in \bigcap\{B' \in \mathcal{P}(B) \mid \Gamma(B') = B'\}$. Since the other inclusion holds trivially, we have $I_\Gamma = \bigcap\{B' \in \mathcal{P}(B) \mid \Gamma(B') = B'\}$. □

Kripke's model-theoretic construction is a fixed point of a monotone operator. However, the specific case we are considering involves an *arithmetically definable* monotone operator. Kripke works within an arithmetical base structure, as the language of arithmetic can be viewed as a simplified model for natural language. The atomic truths in the extension of Kripke's type-free truth predicate correspond to the atomic truths in the standard model $\mathbb{N}$. More complex truths are then defined inductively from these simpler truths, ultimately leading to a (minimal) fixed point where each sentence in the extension of the truth predicate is grounded.[28] The extension (and similarly, the anti-extension, though I will focus on the extension as in section 2.2.2) forms an arithmetically definable set:

Definition 3.9. Let $\Gamma : \mathcal{P}(\mathbb{N}) \to \mathcal{P}(\mathbb{N})$ be an operator. Γ is *arithmetically definable* if there is a $\mathcal{L}_{\mathsf{Tr}}$-formula $\varphi(X, x, \vec{y})$ (all free variables shown), and number parameters $\vec{a}$ such that[29]

$$\Gamma(B) = \{n \in \mathbb{N} \mid \mathbb{N} \models \varphi(B, n, \vec{a})\} \text{ for } \textit{any } B \subseteq \mathbb{N}^m.$$

(I will mostly suppress mentioning parameters.)

Not all arithmetically definable operators are monotone. But if an operator is defined by a *positive* formulae, then it is monotone:

Definition 3.10. Let X be a relation constant. A $\mathcal{L}_{\mathsf{Tr}}$-formula φ is X-positive if X does not occur in φ, φ has the form $X(t_1, ..., t_n)$, or has the form $\psi \vee \chi$, or $\forall x \psi$, where ψ and χ are X-positive.

28 Further discussion on the non-minimal case will follow shortly.

29 As mentioned in section 3.2.1, the satisfaction relation for first-order languages which allow second-order parameters such as $\mathcal{L}_{\mathsf{Tr}}$ (see definition 3.4) can be understood in terms of full or Henkin semantics. Here, I allow B to denote *any* relation on $\mathbb{N}$. So it's best to understand $\models$ in the sense of second-order logic with full semantics. This is in line with Moschovakis's approach to inductive definitions who works with languages that contain relation symbols for each relation on the target set [106, p. 8].

The key point here is that the relation constant X does not appear within the scope of a negation. As a result, the formula φ, which is X-positive, only makes positive relational claims of the form $X(t_1, ..., t_n)$. (This will ultimately be used to express claims such as '$\lfloor\varphi\rfloor$ is in the extension of the truth predicate E', where E appears only positively.) The relationship between *monotone* operators and *positive* formulae can be formally captured by the following lemma:

Lemma 3.11. Let $\varphi(X, y)$ be a $\mathcal{L}_{\mathsf{Tr}}$-formula. If φ is X-positive, the operator induced by φ, Γ_φ, is monotone.

Proof. First, observe that if $B_0 \subseteq B_1$ and $\varphi(B_0, y)$, then also $\varphi(B_1, y)$ (by induction on the complexity of φ). Consequently,

$$\Gamma_\varphi(B_0) = \{n \in \mathbb{N} \mid \mathbb{N} \models \varphi(B_0, n)\} \subseteq \{n \in \mathbb{N} \mid \mathbb{N} \models \varphi(B_1, n)\} = \Gamma_\varphi(B_1),$$

which means that Γ_φ is monotone. □

Operators induced by positive formulae are called *positive operators* or *positive inductive definitions*. If Γ_φ is a positive operator, we simply write I_φ for its fixed point, and I_φ^α for the α^{th} stage of the fixed point.

To summarize, by lemma 3.11, X-positive formulae $\varphi(X, y)$ induce monotone or positive operators (i.e., positive inductive definitions), which reach fixed points after at most $|X|$ steps (as stated in theorem 3.8). As I will demonstrate in the next section, the Kripke construction is a positive inductive definition induced by an arithmetical formula. Consequently, a *Kripke fixed-point* (to be introduced shortly) will be a set of natural numbers, each of which encodes a $\mathcal{L}_{\mathsf{Tr}}$-sentence contained in the extension of the truth predicate.

3.2.3 *Kripke's Truth Theory – Strong Kleene Version*

The central idea behind Kripke's truth theory is a model-theoretic construction of the extension of the truth predicate for an arithmetical language over the standard model of arithmetic. The resulting truth predicate will be *type-free* and a *partial predicate*. The type-free nature of the truth predicate means that, unlike the standard Tarskian truth predicate, iterations of the truth predicate are well-formed. In other words, expressions like $\mathsf{Tr}(\lfloor\mathsf{Tr}(\lfloor\varphi\rfloor)\rfloor)$, which would be ungrammatical in the hierarchical, Tarskian framework, are perfectly well-formed in this context.[30] The fact that the truth predicate is only a partial predicate means that the union of its extension and antiextension forms only a proper subset of the set of sentences of the underlying language. In other words, some sentences—most notably, the liar sentence—are neither in the extension nor the antiextension, but are instead considered *defective* (more on this follows).[31]

30 The key difference is that in the standard Tarskian setting, the truth predicate operates in the metalanguage and only has (codes of) object-language sentences in its extension. Therefore, an expression like $\mathsf{Tr}(\lfloor\mathsf{Tr}(\lfloor\varphi\rfloor)\rfloor)$ would break the boundary between the object language and the metalanguage, making it ungrammatical. For further details, see Halbach [57, Part II].

31 In the following presentation of the Kripke construction, it is theoretically possible for the extension and the antiextension to overlap—that is, some sentences may be classified as both true

The construction proceeds as follows: begin with a relational language of truth in the sense of $\mathcal{L}_{\mathsf{Tr}}$ (section 3.2.1), and an acceptable structure as defined in definition 3.5 that contains a copy of the standard model of arithmetic, $\mathbb{N}$. The goal is to construct a set of (codes of) $\mathcal{L}_{\mathsf{Tr}}$-sentences to serve as the extension of the truth predicate. To start, take the empty set and, in the first step, add all (codes of) atomic sentences from the truth-free fragment of the language that are true in the standard model, as well as all negations of the truth-free fragment of the language that are false in the standard model. We call this set E_1. Thus, E_1 contains all (codes of) sentences of the form $n = n$, and $R(t_1, \ldots, t_m)$ (for all atomic relation symbols) that are true in $\mathbb{N}$, as well as all (codes of) sentences of the form $\neg(n = m)$ and $\neg R(t_1, \ldots, t_m)$ (for all atomic relation symbols) that are false in $\mathbb{N}$.

Next, we add progressively more complex sentences through Strong Kleene satisfaction and the application of the truth predicate. For example, if φ is in E_1, we add sentences like $\neg\neg\varphi$, $\varphi \vee \psi$ (for arbitrary ψ), and $\mathsf{Tr}(\lfloor\varphi\rfloor)$ to obtain E_2.

It is important to note that in order to move from E_n to E_{n+1}, we must refer to E_n, but only positively, meaning that we only add sentences to E_{n+1} that are generated from sentences already in E_n (and not by using a condition of the form $\varphi \notin E_n$). This process is iterated until we reach a point E at which no new sentences are added. At this stage, E is the extension of the truth predicate.

We can formalize the construction using the theory of inductive definitions. The key point is that the Strong Kleene satisfaction predicate can be formalized by a positive formula, which in turn induces a monotone operator. This operator has the set E, as described above, as its fixed-point. To begin, let us define strong Kleene satisfaction:

Definition 3.12. Let $\mathcal{L}^1$ be a first-order language, and $\mathcal{M}$ an $\mathcal{L}^1$-structure. Define strong Kleene satisfaction, in symbols $\models_{sk}$, as follows:[32]

- $\mathcal{M} \models_{sk} a = b$ if and only if $a^{\mathcal{M}} = b^{\mathcal{M}}$,
- $\mathcal{M} \models_{sk} R(t_1, \ldots, t_n)$ if and only if $\langle t_1^{\mathcal{M}}, \ldots, t_n^{\mathcal{M}} \rangle \in R^{\mathcal{M}}$,
- $\mathcal{M} \models_{sk} \neg\neg\psi$ if and only if $\mathcal{M} \models_{sk} \psi$,
- $\mathcal{M} \models_{sk} \psi \vee \chi$ if and only if $\mathcal{M} \models_{sk} \psi$ or $\mathcal{M} \models_{sk} \chi$,
- $\mathcal{M} \models_{sk} \neg(\psi \vee \chi)$ if and only if $\mathcal{M} \models_{sk} \neg\psi$ and $\mathcal{M} \models_{sk} \neg\chi$,
- $\mathcal{M} \models_{sk} \forall x \psi(x)$ if and only if, for every $a \in M$, $\mathcal{M} \models_{sk} \varphi(a)$.

and false; such sentences are called *glutty*. The resulting fixed-point semantics is four-valued, with the truth values 'true', 'false', 'both', and 'neither'. This is a generalization of Kripke's original construction due to Visser [162] and Woodruff [171]. As explained below, a fixed-point is said to be *consistent* if no glutty sentences occur (i.e., if the extension and the antiextension are disjoint). Since I will only deal with consistent fixed-points, I ignore the possibility of a fourth truth value ('both') for the rest of this work.

32 It is important to note that unlike the definition of Tarskian satisfaction, the right-hand side of each clause of definition 3.12 is positive. In particular the clause for negation of Tarskian satisfaction (see definition 3.42) reads "$\mathcal{M} \models \neg\psi$ iff $\mathcal{M} \not\models \psi$", where the right-hand side is negative. This is essentially the reason why we can turn strong Kleene satisfaction into a monotone operator.

Next, I turn definition 3.12 into a positive formula, ensuring that it induces a monotone operator corresponding to strong Kleene satisfaction. As anticipated, I define this operator on the domain of an acceptable base structure $\mathcal{M}$ (definition 3.5). The starting point is a (possibly empty) set E of ($\mathcal{M}$-copies of) natural numbers, which code the sentences that should be included in the extension of the truth predicate:

Definition 3.13. Let $\mathcal{M}$ be an acceptable $\mathcal{L}_{\mathsf{Tr}}$-structure with domain M. Let $E \subseteq N \subseteq M$ (where N is the $\mathcal{M}$-copy of $\mathbb{N}$), and define E^+ as follows: $a \in E^+$ if and only if

- $a \in E$, or
- $a = \lfloor \varphi \rfloor$ for an atomic base $\mathcal{L}_{\in}$-sentence φ and $\mathcal{M} \models \varphi$, or
- $a = \lfloor \neg\neg\psi \rfloor$, $\psi \in \mathsf{Sent}_{\mathsf{Tr}}$ and $\psi \in E$, or
- $a = \lfloor \psi \vee \chi \rfloor$, $\psi, \chi \in \mathsf{Sent}_{\mathsf{Tr}}$ and $\psi \in E$ or $\chi \in E$, or
- $a = \lfloor \neg(\psi \vee \chi) \rfloor$, $\psi, \chi \in \mathsf{Sent}_{\mathsf{Tr}}$ and $\psi \in E$ and $\chi \in E$, or
- $a = \lfloor \forall x \varphi(x) \rfloor$, $\forall x \varphi(x) \in \mathsf{Sent}_{\mathsf{Tr}}$, and for all closed $\mathcal{L}_{\mathsf{Tr}}$-terms t, $\lfloor \varphi(t) \rfloor \in E$, or
- $a = \lfloor \mathsf{Tr}(t) \rfloor$ for a closed $\mathcal{L}_{\mathsf{Tr}}$, $t = \lfloor \psi \rfloor$, $\psi \in \mathsf{Sent}_{\mathsf{Tr}}$, and $\lfloor \psi \rfloor \in E$.

Since the above definition is E-positive, it induces a monotone operator Γ. Let $\xi(a, E)$ abbreviate the right-hand side of definition 3.13 (i.e., everything in the list following the 'if and only if'). Then, starting from $E_0 \subseteq M$, we define

$$\Gamma_\xi(E_0) := \{n \in M \mid \mathcal{M} \models \xi(a, E_0)\},$$

which generates the anticipated sequence of growing subsets of N, indexed by ordinals,

$$E_0, E_1, E_2, \ldots, E_\alpha, \ldots,$$

and eventually reaches a fixed point by theorem 3.8, such that for some ordinal γ, $E_{\gamma+1} = E_\gamma$. We call this fixed point E.

Note that the first clause of definition 3.13 allows for some flexibility. We can either start from the empty set or from a non-empty set. In the first case, the fixed point is called the *minimal fixed point*, and in the latter case, we obtain a *non-minimal fixed point*. This distinction will be important for the contextualist truth theory to be introduced in section 3.2.4.

Let me end this section by saying something about the aforementioned antiextension. What I have described above is actually a variation of the Kripke construction because it only yields the extension. In the original paper, Kripke introduced his truth theory as a way to construct both the extension and the anti-extension of the truth predicate. The construction of the anti-extension is analogous to that of the extension, but starts from atomic base sentences of the truth-free fragment of the language that are *false* in the standard model. In the next steps, the construction proceeds by adding sentences such as $\lfloor \neg\varphi \rfloor$,

where φ is an element in the extension. (Note that when we construct both the extension and the anti-extension simultaneously, we can add negations by reference to the counterpart, i.e., we can add negations to the extension by reference to the anti-extension and *vice versa*.) This ultimately yields a model of the form $\langle \mathcal{M}, \mathsf{E}, \mathsf{A} \rangle$, where the additional A is the anti-extension of the truth predicate. The collection of all sentences that are neither in E nor in A is called the *gap*.

When working with $\langle \mathcal{M}, \mathsf{E}, \mathsf{A} \rangle$, the most natural approach is to use a three-valued logic that assigns to each sentence in E the truth value 1, to each sentence in A the truth value 0, and to all other sentences the intermediate truth value $1/2$. If we work with the minimal fixed-point (i.e., the case where the construction process starts from $\emptyset$), then $\mathsf{E} \cap \mathsf{A} = \emptyset$, meaning that the construction is consistent (since no sentence is both true and false), and $\mathsf{E} \cup \mathsf{A} \subset \mathsf{Sent}^1$, meaning that the truth predicate is a partial predicate. By construction, self-referential sentences such as the liar and the truth-teller ($\tau :\leftrightarrow \mathsf{Tr}(\lfloor \tau \rfloor)$) are neither in the extension nor in the anti-extension.[33] Moreover, the three-valued models validate the naïve truth principles $\mathsf{Tr}-\mathsf{I}$ and $\mathsf{Tr}-\mathsf{E}$ (see section 2.2.2). The reason is that, by construction, $\mathsf{Tr}(\lfloor \varphi \rfloor)$ is in E (A) if and only if φ is in E (A).

I presented only the construction of the extension for several reasons: for the contextualist approach to truth, to be described in section 3.2.4, it is sufficient to work only with the extension of the truth predicate and to construct a sequence of *non-minimal* fixed-points, one for each context. Moreover, the full construction requires additional background on inductive definitions, specifically a *simultaneous inductive definition*[34], which is not necessary if one only intends to construct the extension. However, there is another point to consider: if we use only the extension, we restore classical logic. Essentially, we achieve the same result by starting from the standard Kripke construction $\langle \mathcal{M}, \mathsf{E}, \mathsf{A} \rangle$, and then merging the anti-extension with the gap to obtain $\langle \mathcal{M}, \mathsf{E} \rangle$. This process is sometimes referred to as the closing-off of the model $\langle \mathcal{M}, \mathsf{E}, \mathsf{A} \rangle$. In the closed-off models, $\mathsf{Tr}(\lfloor \varphi \rfloor)$ holds for all sentences that are not in the extension, regardless of whether they would be in the anti-extension or the gap. The closed-off models $\langle \mathcal{M}, \mathsf{E} \rangle$ also correspond to the intended models of Feferman's fully classical axiomatization of Kripke's truth theory, KF.[35]

3.2.4 *Contextualism à la Glanzberg à la Rossi*

The form of contextualism used here traces back to Charles Parsons [118] and was later expanded by Michael Glanzberg [50, 47]. This presentation follows Rossi [144], who simplifies Glanzberg's approach.[36]

33 In the minimal fixed-point, all sentences in the extension and anti-extension are ultimately grounded in (arithmetic) facts. Self-referential sentences do not bottom out in facts and are therefore *un*grounded. According to Kripke, this is a kind of defective form of language use. See also footnote 52 in chapter 2.

34 See Moschovakis [106, Lemma 1.C.I].

35 See section 3.2.5 as well as Feferman [33], Cantini [19], and Halbach and Horsten [58] and Halbach [57, ch. 15].

36 The main difference is that Rossi introduces a new truth predicate for each context, while Glanzberg uses a single one for all contexts. Rossi notes that this is merely a technical simplifica-

Parsons-Glanzberg Contextualism models contexts through a sequence of Kripkean fixed points. In context 0, the minimal fixed point E_0 is constructed over an arithmetical base structure $\mathcal{M}_0$ as outlined in section 3.2.3.[37] The context 0 liar is not in the c_0-extension of the truth predicate (i.e., $\lambda_0 \notin \mathsf{E}_0$), and therefore is false in c_0. This simply means that λ_0 does not express a true proposition in c_0.[38] The context 0 liar triggers a shift to the subsequent context c_1, in which a non-minimal fixed point E_1 is constructed. E_1 builds on the c_0-extension of the truth predicate, so that λ_0 is true in c_1. This is because λ_0 now states in c_1 that λ_0 doesn't express a true proposition in c_0, which is true.

Technically, we use a sequence of languages, each with a unique truth predicate Tr_α for expressing *truth-in-*c_α. We also have a sequence of contextualized liar sentences λ_α. The construction begins by building the minimal fixed point E_0 over the acceptable structure $\mathcal{M}_0$ as the extension of Tr_0. This yields the structure $\langle \mathcal{M}_0, \mathsf{E}_0 \rangle$, which we use as a model for *truth-in-*c_0. Notably, $\lambda_0 \notin \mathsf{E}_0$, so $\langle \mathcal{M}_0, \mathsf{E}_0 \rangle \not\models \mathsf{Tr}_0(\lfloor \lambda_0 \rfloor)$. Classical logic implies that $\langle \mathcal{M}_0, \mathsf{E}_0 \rangle \models \neg \mathsf{Tr}_0(\lfloor \lambda_0 \rfloor)$, which in turn leads to $\langle \mathcal{M}_0, \mathsf{E}_0 \rangle \models \lambda_0$.

For context 1, consider the closing-off of c_0 and recall that it contains all sentences validated in c_0, i.e., $\mathsf{ClOff}_0 := \{\varphi \mid \langle \mathcal{M}_0, \mathsf{E}_0 \rangle \models \varphi\}$. To construct the extension of Tr_1, take the closing-off of c_0 as the starting point and construct a new non-minimal fixed point E_1. Since $\mathsf{ClOff}_0 \subseteq \mathsf{E}_1$, λ_0 belongs to Tr_1, meaning $\langle \mathcal{M}_1, \mathsf{E}_1 \rangle \models \mathsf{Tr}_1(\lfloor \lambda_0 \rfloor)$. Thus, in c_1, it is true that λ_0 does not express a true proposition in c_0. Of course, c_1 comes with its own liar sentence λ_1, which does not express a true proposition in c_1 (i.e., $\langle \mathcal{M}_1, \mathsf{E}_1 \rangle \models \lambda_1$), and from the perspective of c_2, we can say that *it is true in* c_2 that λ_1 does not express a true proposition in c_1 (i.e., $\langle \mathcal{M}_2, \mathsf{E}_2 \rangle \models \mathsf{Tr}_2(\lfloor \lambda_1 \rfloor)$), and so on.

In order to model contextualism, let me first define the sequence of object languages that we use to model contextualism:

Definition 3.14. Let $\mathfrak{L} := \mathcal{L}^1_1, \mathcal{L}^1_2, \ldots, \mathcal{L}^1_\alpha, \ldots$ be a sequence of countable first-order languages satisfying:

(a) Each $\mathcal{L}^1_\alpha$ includes $\in$ but has no function symbols.

(b) Each $\mathcal{L}^1_\alpha$ has a fresh unary truth predicate Tr_α for context α.

(c) $\mathcal{L}^1_\alpha$ and $\mathcal{L}^1_{\alpha+1}$ share terms except for $\mathsf{Tr}_{\alpha+1}$, present only in $\mathcal{L}^1_{\alpha+1}$.

(d) Each $\mathcal{L}^1_\alpha$ has at least one acceptable structure $\mathcal{M}^\alpha$.

(e) For each structure $\mathcal{M}^\alpha$ and expression e, there exists a term $\lfloor e \rfloor$ denoting e's code in $\mathcal{M}^\alpha$.

(f) For any open formula $\varphi(x)$, there is a term t_φ such that $(t_\varphi)^{\mathcal{M}^\alpha} = (\lfloor \varphi(t_\varphi/x) \rfloor)^{\mathcal{M}^\alpha}$.

---

tion, meaning that both the contextualist and bicontextualist frameworks presented below can be expressed using only a single truth predicate (see Rossi [144, Sec. 5.1]).

37 In section 3.2.3, E_α represents the stages of the inductive definition, while E denotes the fixed point. Here, E_α indicates the (possibly non-minimal) fixed point of context α.

38 See section 2.2.2 for the rephrasing of the liar sentence as: "it is not the case that the context 0 liar expresses a true proposition in context 0."

Condition (a) ensures that each language contains $\in$ to develop mathematical notions in their set-theoretic representation, and only relation symbols to simplify the following definitions. Conditions (b) and (c) imply that $\mathcal{L}^1_\alpha \subseteq \mathcal{L}^1_{\alpha+1}$, with each context $c_{\alpha+1}$ and language $\mathcal{L}^1_{\alpha+1}$ adding a fresh unary predicate expressing *truth-in-*$c_{\alpha+1}$. Condition (d) ensures that each language $\mathcal{L}^1_\alpha$ has an acceptable, set-sized structure $\mathcal{M}^1_\alpha$ in the sense of section 3.2.1. Requirements (e) and (f) ensure that each language has sufficiently many terms for coding. By (f), the strong diagonalization lemma is provable.

Let me now formally define the closing-off construction:

Definition 3.15. Let $\mathcal{M}_0$ be an acceptable base structure, E_0 the minimal Kripkean fixed point over $\mathcal{M}_0$, and I_ξ the monotone operator induced by strong Kleene satisfaction (definition 3.12). Define the closing-off for $\mathcal{L}^1_0$ as:

$$\mathsf{ClOff}_0 := \{\varphi \in \mathsf{Sent}^1_0 \mid \langle \mathcal{M}_0, \mathsf{E}_0\rangle \models \varphi\}.$$

For $\mathcal{M}_\alpha$, the closing-off for $\mathcal{L}^1_\alpha$ is:

$$\mathsf{ClOff}_\alpha := \{\varphi \in \mathsf{Sent}^1_\alpha \mid \langle \bigcup_{\beta\leq\alpha} \mathcal{M}_\beta, I_\xi(\bigcup_{\beta<\alpha} \mathsf{ClOff}_\beta)\rangle \models \varphi\}.$$

First, recall that Γ_ξ denotes the monotone operator induced by the strong Kleene satisfaction schema (see definition 3.12), and that I_ξ denotes its fixed point. I also use E with ordinal indices for fixed points. In the above definition, the first line uses the minimal Kripkean fixed point to define the closing-off of c_0. $\langle \mathcal{M}_0, \mathsf{E}_0\rangle$ is the model for the initial context, where $I_\xi(\emptyset) = \mathsf{E}_0$ (i.e., E_0 is the minimal fixed point of c_0).

In c_1, we use the closing-off of the minimal Kripkean fixed point from c_0 to construct the non-minimal Kripkean fixed point $I_\xi(\mathsf{ClOff}_0)$. Thus, in c_1, the construction starts with ClOff_0 rather than with the empty set. Iterating the monotone operator Γ_ξ yields the stepwise construction of the non-minimal fixed point, where this time we close with Tr_1. Therefore, we obtain the sequence

$$\mathsf{ClOff}_0 \subseteq \Gamma_\xi(\mathsf{ClOff}_0) \subseteq \Gamma_\xi(\mathsf{ClOff}_0 \cup \Gamma_\xi(\mathsf{ClOff}_0)) \subseteq \cdots \subseteq \mathsf{E}_1 = \Gamma_\xi(\mathsf{E}_1),$$

and use its fixed point $I_\xi(\mathsf{ClOff}_0) = \mathsf{E}_1$ as the extension of Tr_1, with $\langle \mathcal{M}_0 \cup \mathcal{M}_1, I_\xi(\mathsf{ClOff}_0)\rangle$ serving as the model for c_1.

In the next step, we close-off again to get ClOff_1 which we use in c_2. This time, this requires to apply Γ_ξ to the collection of the previous closing-offs, i.e., we get the sequence

$$\begin{aligned}\mathsf{ClOff}_0 \cup \mathsf{ClOff}_1 &\subseteq \Gamma_\xi(\mathsf{ClOff}_0 \cup \mathsf{ClOff}_1) \subseteq \\ &\Gamma_\xi(\mathsf{ClOff}_0 \cup \mathsf{ClOff}_1 \cup \Gamma_\xi(\mathsf{ClOff}_0 \cup \mathsf{ClOff}_1)) \subseteq \ldots \subseteq \mathsf{E}_2 = \Gamma_\xi(\mathsf{E}_2).\end{aligned}$$

So we get $\langle \mathcal{M}_0 \cup \mathcal{M}_1 \cup \mathcal{M}_2, I_\xi(\mathsf{ClOff}_0 \cup \mathsf{ClOff}_1)\rangle$ for c_2, and so on. Definition 3.15 yields a sequence of closing-offs for each context. For contexts c_α and $c_{\alpha+1}$, we have:

$$\langle \bigcup_{\beta\leq\alpha} \mathcal{M}_\beta, \Gamma_\xi(\bigcup_{\beta<\alpha} \mathsf{ClOff}_\beta)\rangle \not\models \mathsf{Tr}_\alpha(\lfloor \lambda_\alpha \rfloor), \tag{3.2}$$

$$\langle \bigcup_{\beta\leq\alpha} \mathcal{M}_\beta, \Gamma_\xi(\bigcup_{\beta<\alpha} \mathsf{ClOff}_\beta)\rangle \models \neg \mathsf{Tr}_\alpha(\lfloor \lambda_\alpha \rfloor), \tag{3.3}$$

$$\langle \bigcup_{\beta\leq\alpha} \mathcal{M}_\beta, \Gamma_\xi(\bigcup_{\beta<\alpha} \mathsf{ClOff}_\beta)\rangle \models \lambda_\alpha, \tag{3.4}$$

$$\langle \bigcup_{\beta\leq\alpha+1} \mathcal{M}_\beta, \Gamma_\xi(\bigcup_{\beta<\alpha+1} \mathsf{ClOff}_\beta)\rangle \models \mathsf{Tr}_{\alpha+1}(\lfloor \lambda_\alpha \rfloor). \tag{3.5}$$

Since $\lambda_\alpha \notin \Gamma_\xi(\bigcup_{\beta<\alpha} \mathsf{ClOff}_\beta)$, which means that λ_α is not true in c_α (3.2). Since $\models$ is classical, 3.3 follows, and by definition of λ_α, also 3.4. This means that λ_α does not express a true proposition in c_α. In the subsequent context $c_{\alpha+1}$, however, it is true (i.e., *true-in-*$c_{\alpha+1}$) that λ_α does not express a true proposition on c_α, i.e., 3.5.

The contextualist truth theory employed here is not a naïve truth theory. Instead, it uses *context parameters* (indexed by ordinals) to stratify the application of the truth predicate. This stratification plays a crucial role in preserving consistency.

It is also worth emphasizing that the contextualist truth predicates Tr_α is, like in Kripke's original fixed-point construction, are *partial predicates*. That is, for some sentences φ, neither $\mathsf{Tr}_\alpha(\lfloor\varphi\rfloor)$ nor $\neg\mathsf{Tr}_\alpha(\lfloor\varphi\rfloor)$ holds at stage α. This partiality is a central feature of the theory: it blocks the application of the Law of the Excluded Middle in certain cases and plays a key role in avoiding paradox.

Formally, the theory allows for *unrestricted truth-elimination*:

$$\mathsf{Tr}_\alpha(\lfloor\varphi\rfloor) \models \varphi,$$

meaning that if a sentence is declared true at level α, its content φ is semantically entailed in that same context. However, *truth-introduction* generally requires a context shift. Specifically, if a sentence φ is determined to be in the closing off at level α, then its truth-ascription is only available at the next level:

$$\text{if } \varphi \in \mathsf{ClOff}_\alpha \text{ then } \mathsf{Tr}_{\alpha+1}(\lfloor\varphi\rfloor).$$

The two principles are collected in the following lemma:

Lemma 3.16. Let Tr_α, $\mathsf{Tr}_{\alpha+1}$ be contextualized truth predicates, $\mathsf{ClOff}\alpha$ be the closing-off of c_α, and $\varphi \in \mathcal{L}^1_\alpha$. Then

1. $\mathsf{Tr}_\alpha(\lfloor\varphi\rfloor) \models \varphi$,
2. if $\varphi \in \mathsf{ClOff}_\alpha$ then $\mathsf{Tr}_{\alpha+1}(\lfloor\varphi\rfloor)$.

Proof.

1. This follows from the semantic clause of the inductive fixed-point construction for the truth predicate. If $\mathsf{Tr}_\alpha(\lfloor\varphi\rfloor)$ holds, then φ belongs to the closing-off of c_α, and is thus semantically valid.

2. If φ is in the closing-off at stage α, then the next stage adds the truth-ascription because the closing-off construction for $c_{\alpha+1}$ takes ClOff_α as the starting point.

□

This asymmetry blocks the derivation of paradoxes in the contextualist truth theory. The resulting system is thus *non-naïve*, and *consistent*. For these reasons, the bicontextualist framework presented here should not be understood as relying on an inconsistent truth theory.

3.2.5 *A Short Note on Axiomatization*

Even though this work mostly concerns a model-theoretic approach to truth, I shall also mention the axiomatic side. Kripke's truth theory has famously been axiomatized by Feferman in his "Reflecting on Incompleteness" [33], and independently by Cantini in his "Notes on Formal Theories of Truth" [19]. The resulting theory, known as KF (for Kripke-Feferman), axiomatizes a type-free truth predicate, based on the strong Kleene version of Kripke's ideas from his "Outline of a Theory of Truth" [79]. The similarity with the fixed-point construction can already be seen from the fact that in the definition, Tr appears only positively. This establishes a connection between the KF axioms and the inductive definition of the Kripkean fixed point as described in section 3.2.3, and vice versa.

To present the theory, I need some further notational conventions.[39] I will use dot notation to represent primitive recursive functions that operate on codes of formulae. For instance, let x be the code of a formula φ. Then, $\dot{\neg}x$ represents the primitive recursive function that takes as input the code of φ and returns the code of the formula $\neg\varphi$, and similarly for $x \dot{=} y$, $x \dot{\wedge} y$, $\dot{\forall} v y$, etc. Furthermore, I use the symbol $\mathsf{val}(x)$ for a valuation function, which, when applied to a closed term, outputs the numeral of the number that it denotes. For example, $\mathsf{val}(s) = t$ expresses that the value of the term s is t (taken to be a numeral), and $\mathsf{val}(s) = \mathsf{val}(t)$ expresses that s and t have the same value.[40] Finally, if x is the code of a formula $\varphi(v)$ in which v occurs free, $x(t/v)$ represents the code of the formula in which v is replaced with t. Here is the theory:

Definition 3.17. The theory KF is given by all the axioms of *Peano Arithmetic* with induction in the language containing the truth predicate, together with the following axioms:

KF1 $\forall s \forall t(\mathsf{Tr}(s \dot{=} t) \leftrightarrow \mathsf{val}(s) = \mathsf{val}(t))$

KF2 $\forall s \forall t(\mathsf{Tr}(\dot{\neg} s \dot{=} t) \leftrightarrow \mathsf{val}(s) \neq \mathsf{val}(t))$

KF3 $\forall x(\mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}(x) \rightarrow (\mathsf{Tr}(\dot{\neg}\dot{\neg} x) \leftrightarrow \mathsf{Tr}(x)))$

39 See, for example, Halbach [57, ch. 5].

40 Note that the language cannot contain a symbol for val. Nevertheless, the function that maps a specific term to the numeral of its value can be expressed in the language, so I follow the usual notational conventions and use val in the following.

KF4 $\forall x \forall y(\mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}(x \dot{\wedge} y) \to (\mathsf{Tr}(x \dot{\wedge} y) \leftrightarrow \mathsf{Tr}(x) \wedge \mathsf{Tr}(y)))$

KF5 $\forall x \forall y(\mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}(x \dot{\wedge} y) \to (\mathsf{Tr}(\dot{\neg}(x \dot{\wedge} y)) \leftrightarrow \mathsf{Tr}(\dot{\neg} x) \vee \mathsf{Tr}(\dot{\neg} y)))$

KF6 $\forall v \forall x(\mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}(\dot{\forall} v x) \to (\mathsf{Tr}(\dot{\forall} v x) \leftrightarrow \forall t \mathsf{Tr}(x(t/v))))$

KF7 $\forall v \forall x(\mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}(\dot{\forall} v x) \to (\mathsf{Tr}(\dot{\neg} \dot{\forall} v x) \leftrightarrow \exists t \mathsf{Tr}(\dot{\neg} x(t/v))))$

KF8 $\forall t(\mathsf{Tr}(\dot{\mathsf{Tr}}(t)) \leftrightarrow \mathsf{Tr}(\mathsf{val}(t)))$

KF9 $\forall t(\mathsf{Tr}(\dot{\neg} \dot{\mathsf{Tr}}(t)) \leftrightarrow (\mathsf{Tr}(\dot{\neg} \mathsf{val}(t)) \vee \neg \mathsf{Sent}_{\mathcal{L}_{\mathsf{Tr}}}(\mathsf{val}(t))))$

Many have advocated for KF.[41] Halbach and Horsten have argued that closed-off models, as defined in definition 3.15, "are the intended models for the Kripke-Feferman theory KF" [58, p. 681], in the sense that if $\mathsf{KF} \vdash \varphi$, then φ holds in all closed-off models, which makes it sound with respect to those models. Moreover, since $\mathsf{KF} \vdash \lambda$ and $\mathsf{KF} \vdash \neg \mathsf{Tr}(\lfloor \lambda \rfloor)$, the liar sentence is treated in KF just as it is in the closed-off models.[42] As we will see in chapter 4, bicontextualism validates KF.

3.3 FIRST- AND SECOND-ORDER SET THEORY

The main aim of this section is to highlight the difference between first- and second-order set theory. In particular, I will show that standard first-order axiomatizations (ZFC) are not powerful enough to pick out only standard models with structurally nice features. In contrast, second-order axiomatizations (ZFC_2) are powerful enough to characterize, up to *quasi-categoricity*, structurally nice models. More precisely, a model is nice when it's isomorphic to a standard model, i.e., a strongly inaccessible rank initial segment of the cumulative hierarchy. Moreover, the theory is quasi-categorical if it all its models are initial segment of the cumulative hierarchy without fixing their height. All these notions will be explained and made more precise within this section.

The first step is to give an exact definition of the theories involved:

Definition 3.18. The theory ZFC is the first-order $\mathcal{L}^1_{\in}$-theory that comprises the following seven axioms and two axiom schemata:[43]

EXTENSIONALITY $\forall x \forall y(\forall z(z \in x \leftrightarrow z \in y) \to x = y)$.

PAIRING $\forall x \forall y \exists z(z = \{x, y\})$.

UNION $\forall x \exists y(y = \bigcup x)$.

POWERSET $\forall x \exists y(y = \mathcal{P}(x))$.

INFINITY $\exists x(\emptyset \in x \wedge \forall y(y \in x \to y \cup \{y\} \in x))$.

FOUNDATION (REGULARITY) $\forall x(\emptyset \neq x \to \exists y(y \in x \wedge y \cap x = \emptyset))$.

41 See McGee [102], Rossi [144], Feferman [33], Halbach and Horsten [58], and Reinhardt [138].

42 See Halbach and Horsten [58], Proposition 3 and Remark 5. Compare also eqs. (3.4) and (3.5) from section 3.2.4.

43 I use the following abbreviations: $\bigcup x = \{y | y \in x\}$, $\mathcal{P}(x) = \{y | y \subseteq x\}$, $y \cap x = \{z | z \in y \wedge z \in x\}$.

CHOICE $\forall x(\emptyset \notin x \to \exists f : x \to \bigcup x \wedge \forall a \in x(f(a) \in a))$.

SEPARATION SCHEMA Let $\varphi(x, \vec{p})$ be an $\mathcal{L}_\in$-formula with all free variables displayed, then the following is a axiom:

$$\forall z \forall \vec{p} \exists y \forall x (x \in y \leftrightarrow x \in z \wedge \varphi(x, \vec{p})).$$

REPLACEMENT SCHEMA Let $\varphi(x, y, \vec{p})$ be an $\mathcal{L}_\in$-formula with all free variables displayed, then the following is a axiom:

$$\forall z \forall \vec{p} (\forall x (x \in z \to \exists! y \varphi(x, y, \vec{p})) \to \exists a \forall x (x \in z \to \exists y (y \in a \wedge \varphi(x, y, \vec{p})))).$$

Since ZFC is a first-order theory, we can prove the existence of non-standard ZFC-models via the compactness and Löwenheim-Skolem theorems of first-order logic (seesection 3.3.3). This fact will be crucial in the rest of this section and in chapter 5. A way to overcome this limitation is by using a second-order theory:

Definition 3.19. The theory ZFC_2 is formulated in the *second-order* language $\mathcal{L}^2_\in$, and it results from ZFC by replacing the schemata of separation and replacement by the following axioms:

SEPARATION AXIOM $\forall F \forall z \exists y \forall x (x \in y \leftrightarrow x \in z \wedge F(x))$.

REPLACEMENT AXIOM

$$\forall F \forall z \forall \vec{p} (\forall x (x \in z \to \exists! y F(x, y, \vec{p})) \to \exists a \forall x (x \in z \to \exists y (y \in a \wedge F(x, y, \vec{p})))).$$

Unlike ZFC, ZFC_2 is formulated in second-order logic which lacks the metalogical properties needed to prove the existence of non-standard models. In what follows, I will spell this out in more detail by proving the following theorems:

Theorem 3.20. Let $\mathcal{M}$ be a transitive structure such that $\mathcal{M} \models \mathsf{ZFC}$. Then there exists a countable transitive model $\mathcal{N} \equiv \mathcal{M}$.

Theorem 3.21. Let $\mathcal{M}, \mathcal{N}$ be models of ZFC_2. Then exactly one of the following obtains:

1. $\mathcal{M} \cong \mathcal{N}$
2. $\mathcal{M} \cong V_\kappa^{\mathcal{N}}$, for some κ which $\mathcal{N}$ thinks is an inaccessible cardinal.
3. $\mathcal{N} \cong V_\kappa^{\mathcal{M}}$, for some κ which $\mathcal{M}$ thinks is an inaccessible cardinal

A full presentation of all the material is beyond the scope of this work, so I will focus on the most important points for each of the theorems. As the main issue in chapter 5 is on structurally nice models, I'll focus on theorem 3.21. In combination, theorems 3.20 and 3.21 imply the existence of ZFC-models which lack the structural *niceness*, and rule out the existence of ZFC_2-models that are not nice in this sense.[44]

44 For the most part, this section follows Button and Walsh [18, ch. 8.A]. The part on absoluteness follows Devlin [27, p. 24 ff.].

3.3.1 *Some Background Notions*

In section 2.2.1, I introduced the cumulative hierarchy as the ontological (model-theoretic) counterpart to the iterative conception of sets. Recall that the cumulative hierarchy is defined by transfinite recursion as follows:

$$V_0 := \emptyset, \quad V_{\alpha+1} := \mathcal{P}(V_\alpha), \quad V_\gamma := \bigcup_{\alpha<\gamma} V_\alpha \quad \text{(for limit ordinals } \gamma\text{)}.$$

The cumulative hierarchy, V, is defined as the union of all these ranks:

$$\mathsf{V} := \bigcup_{\alpha\in\mathrm{Ord}} V_\alpha.$$

An important aspect of the cumulative hierarchy is that its ranks are *transitive*, i.e., if $x \in y \in V_\alpha$, then $x \in V_\alpha$. Moreover, if $\alpha \leq \beta$, then $V_\alpha \subseteq V_\beta$, and $\alpha \subseteq V_\alpha$ for all ordinals α.[45] Note that this final claim implies that for all ordinals α, $\mathrm{rank}(\alpha) = \alpha$.[46]

The cumulative hierarchy contains the standard models of ZFC. In the next section, I give a precise characterization of these standard models, in particular of the closure properties (closure under set-of operations), they have to satisfy. However, before I introduce inaccessible cardinals, let me first state some results that will be important here and also later on. Let me begin with *Mostowski Collapse Theorem* ([107]) which shows that we're justified to focus on transitive structure like the ones contained in V:

Theorem 3.22 (Mostowski Collapse Theorem)**.** Any well-founded, extensional structure is isomorphic to a transitive structure.

Proof. Let $\mathcal{M}$ be well-founded and extensional. More precisely, let $\mathcal{M}$ be a structure such that $\mathcal{M} \models$ EXTENSIONALITY, and such that there are no infinite descending $\in$-chains in $\mathcal{M}$ (i.e., there is no sequence of sets a_n such that for all $n \in \mathbb{N}$, $\mathcal{M} \models a_{n+1} \in a_n$). We define a rank function for $\mathcal{M}$ as follows:

$$\mathrm{rank}^{\mathcal{M}}(m) = \begin{cases} 0, & \text{if } \mathcal{M} \models \forall x\ x \notin m \\ \sup\{\mathrm{rank}^{\mathcal{M}}(n)+1 \mid \mathcal{M} \models n \in m\}, & \text{otherwise.} \end{cases}$$

Since $\mathcal{M}$ is extensional and well-founded, $\mathrm{rank}^{\mathcal{M}}$ is well-defined. We now define a *collapse function*, h, on $\mathrm{rank}^{\mathcal{M}}$, such that $\mathrm{rank}^{\mathcal{M}}(x) = \mathrm{rank}(h(x))$ by transfinite recursion as follows:

$$h(m) = \begin{cases} \emptyset, & \text{if } \mathrm{rank}^{\mathcal{M}}(m) = 0 \\ \{h(n) \mid \mathcal{M} \models n \in m\}, & \text{otherwise.} \end{cases}$$

45 Proof sketch: $V_0 = \emptyset$ is vacuously transitive. Now assume that for all $\beta < \alpha$, V_β is transitive. Let $x \in y \in V_\alpha$. We want to show that $x \in V_\alpha$. If $\alpha = \beta + 1$, $y \in V_\alpha$ implies $y \subseteq V_\beta$, and so $x \in V_\beta$. Since V_β is transitive by IH, $x \subseteq V_\beta$, so $x \in \mathcal{P}(V_\beta) = V_{\beta+1}$. If α is limit, $y \in V_\alpha$ implies that $y \in V_\gamma$ for some $\gamma < \alpha$. Then, since V_γ is transitive by IH, $y \subseteq V_\gamma$, and so $x \in V_\gamma \subseteq \bigcup_{\gamma<\alpha} V_\gamma = V_\alpha$. The transitivity of the V_α's implies that $V_\alpha \subseteq V_\beta$ for $\alpha < \beta$. Finally, $0 = \emptyset \subseteq V_0$. Now assume that $\beta \subseteq V_\beta$ for all $\beta < \alpha$. We want to show that $\alpha \subseteq V_\alpha$. If $\alpha = \gamma + 1$, we have to show that $\alpha = \gamma \cup \{\gamma\} = \{0, 1, ..., \gamma\} \subseteq V_\alpha$. By IH, $\gamma \subseteq V_\gamma$, and so $\gamma \in \mathcal{P}(V_\gamma) = V_\alpha$. Since V_α is transitive, $\gamma \subseteq V_\alpha$, and so $\{0, 1, ..., \gamma\} \subseteq V_\alpha$. Let α be a limit ordinal. Then, $V_\alpha = \bigcup_{\gamma<\alpha} V_\gamma$. By IH, $\gamma \subseteq V_\gamma$, so $\gamma \in V_\alpha$, and therefore $\{\gamma | \gamma < \alpha\} = \alpha \subseteq V_\alpha$.

46 The rank-function is defined as $\mathrm{rank}(x) =$ the least α such that $x \in V_{\alpha+1}$ (see [75, p. 64 f.]). Consequently, $V_{\alpha+1}$ is the least stage of the cumulative hierarchy that contains α as a set.

With the collapse function, we can define a model $\mathcal{N} = \langle N, \in^{\mathcal{N}} \rangle$, called the *transitive collapse*, or the *Mostowski collapse* of $\mathcal{M}$, where

$$N = \{h(x) \mid x \in \mathcal{M}\}$$
$$\in^{\mathcal{N}} = \{(n, m) \in N \times N \mid n \in m\}.$$

$\mathcal{N}$ is transitive: let $n \in m \in N$. Then, there is an $m' = h(m') = \{h(n') \mid \mathcal{M} \models n' \in m'\}$. But then, $n = h(n')$ for a n' such that $\mathcal{M} \models n' \in m'$, and so $n \in N$.

It remains to be shown that h is an isomorphism. Since h is surjective by definition, it suffices to show that h is injective and preserves structure. Injectivity follows by induction on $\text{rank}^{\mathcal{M}}$: for the base case, let $m, n \in M$ be such that $\text{rank}^{\mathcal{M}}(m) = \text{rank}^{\mathcal{M}}(n) = 0$. Then, $h(m) = \{m' \in M \mid m' \in m\} = \emptyset = \{n' \in M \mid n' \in n\} = h(n)$. Now assume that if $\max\{\text{rank}^{\mathcal{M}}(x), \text{rank}^{\mathcal{M}}(y)\} < \alpha$, then $h(x) = h(y)$. Let $n, m \in M$ be such that $\text{rank}^{\mathcal{M}}(n) = \text{rank}^{\mathcal{M}}(m) = \alpha$ and assume $h(n) = h(m)$. If $\mathcal{M} \models d \in n$, then $h(d) \in h(n) = h(m)$, so for some $e \in M$, $\mathcal{M} \models e \in m$ and $h(d) = h(e)$, and by induction hypothesis, $d = e$, so $\mathcal{M} \models d \in m$. The same holds for the other direction, i.e., if $\mathcal{M} \models e \in n$, then $\mathcal{M} \models e \in m$. Since $\mathcal{M}$ is extensional, $n = m$.

Finally, h preserves structure. If $\mathcal{M} \models n \in m$, then, by definition of h, $h(n) \in h(m)$. On the other hand, if $h(n) \in h(m)$, again by definition of h, $h(n) = h(n')$ for some n' such that $\mathcal{M} \models n' \in m$. But then $n' = n$ since h is injective, and so $\mathcal{M} \models n \in m$. □

Next, the notion of *absoluteness* will be important here and in chapter 5. Absoluteness is a property of $\mathcal{L}^1_\in$-formulae. Roughly, a formula is absolute for a certain class of structures if it has the same truth value in each structure of the class. Absoluteness results typically depend on the complexity of formulae:

Definition 3.23. A formula of $\mathcal{L}^1_\in$ is

- Δ_0 if it has no quantifiers, or it is of the form $\varphi \wedge \psi$, $\varphi \vee \psi$, $\neg\varphi$, or $\varphi \to \psi$, where φ and ψ are Δ_0, or it is of the form $(\exists x \in y)\varphi$ or $(\forall x \in y)\varphi$ where φ is Δ_0;
- Σ_1 if it is of the form $\exists x\varphi$ where φ is Δ_0;
- Π_1 if it is of the form $\forall x\varphi$ where φ is Δ_0.

Moreover, we need to relativize formulae to certain (class-sized) structures in order to express their absoluteness for those structures. If φ is an $\mathcal{L}^1_\in$-formula and $\mathcal{M}$ an $\mathcal{L}_\in$-model, let $\varphi^{\mathcal{M}}$ result from φ by relativizing all quantifiers in φ to M. More precisely, if φ is atomic, then $\varphi^{\mathcal{M}} = \varphi$. If φ is of the form $\varphi \wedge \psi$, $\varphi \vee \psi$, $\neg\varphi$, or $\varphi \to \psi$, then $\varphi^{\mathcal{M}}$ is of the form $\varphi^{\mathcal{M}} \wedge \psi^{\mathcal{M}}$, $\varphi^{\mathcal{M}} \vee \psi^{\mathcal{M}}$, $\neg\varphi^{\mathcal{M}}$, or $\varphi^{\mathcal{M}} \to \psi^{\mathcal{M}}$. The important case is the quantifier case. Let φ be of the form $\exists x\psi(x)$. Since we want to relativize φ to M, but M typically is a class, we must restrict the quantifier to the formula that defines M. So, let $M := \{x \mid \chi(x)\}$, then $\varphi^{\mathcal{M}} = (\exists x\psi(x))^{\mathcal{M}}$ is defined to be $\exists x(\chi(x) \wedge \psi^{\mathcal{M}}(x))$.

With the definition of the relativization, we can give an exact definition of the absoluteness property:

Definition 3.24. Let $\mathcal{M}$ be a transitive model. An $\mathcal{L}^1_\in$-formula is

- *upwards absolute* iff for all $\vec{x}$, $\varphi^{\mathcal{M}}(\vec{x}) \to \varphi(\vec{x})$;
- *downwards absolute* iff for all $\vec{x}$, $\varphi(\vec{x}) \to \varphi^{\mathcal{M}}(\vec{x})$;
- *absolute* iff it's upwards and downwards absolute, i.e., iff for all $\vec{x}$, $\varphi(\vec{x}) \leftrightarrow \varphi^{\mathcal{M}}(\vec{x})$.

In this work, I'm exclusively interested in transitive structures (that are models of ZFC) in the sense of the cumulative hierarchy. So, for the absoluteness theorem, it is of great importance to have this notion of structure in mind. Δ_0-formulae (which only contain restricted quantifiers) will have the same truth value throughout the whole hierarchy. The reason is that we only consider rank-initial segments of V that are ZFC-models. And so either every such model validates the defining formula that we use for the restriction (the formula $\chi(x)$ used above), or none of these models does. Hence, Δ_0-formulae will be absolute for transitive structures. Regarding upwards absoluteness, any set, once it enters the cumulative hierarchy, will remain there, and so it will be contained in any later level. This means that existential claims (Σ_1-formulae) will be true once their witness is constructed, and so they keep their truth value as we move upwards in the hierarchy. Regarding downwards absoluteness, universal claims (Π_1-formulae) that are true in the whole hierarchy remain true no matter how far down we go. This is because any lower rank of V, from that point of view, is like a quantifier domain restriction. And so any universal property that holds true unrestrictedly, also holds true when we restrict the quantifiers to any lower rank.

Theorem 3.25. Let $\mathcal{M}$ be a transitive model and φ an $\mathcal{L}^1_\in$-formula.

1. If φ is Δ_0, then φ is absolute for $\mathcal{M}$;
2. If φ is Σ_1, then φ is upwards absolute;
3. If φ is Π_1, then φ is downwards absolute.

Proof.

1. If φ contains no quantifiers (i.e., is atomic, or of φ is of the form $\psi \wedge \chi$, $\psi \vee \chi$, $\neg\psi$, or $\psi \to \chi$ where both ψ and χ contain no quantifiers), then $\varphi = \varphi^{\mathcal{M}}$.
   Now let φ be $(\exists x \in y)\psi(x)$ and let $y \in M$. Assume, for induction, that $\psi^{\mathcal{M}}(x) \leftrightarrow \psi(x)$. If $\varphi(x)^{\mathcal{M}}$, we get $[(\exists x \in y)(\psi(x,y))]^{\mathcal{M}}$ for some $x \in M$ such that $x \in y$. But then, we get $(\exists x \in y)\psi^{\mathcal{M}}(x)$ and by induction hypothesis, $(\exists x \in y)\psi(x)$, i.e., φ. Conversely, assume $(\exists x \in y)\psi(x)$. Then, for some $x \in y$, $\psi(x)$. Since $\mathcal{M}$ is transitive, $y \in M$ implies that $x \in M$, and so by the induction hypothesis, $\psi^{\mathcal{M}}(x)$, and so we get $[(\exists x \in y)(\psi(x,y))]^{\mathcal{M}}$, i.e., $\varphi^{\mathcal{M}}$. (The case for $(\forall x \in y)\psi(x)$ is analogous.)
2. Let φ be $\exists u\psi(u)$, ψ be Δ_0, and assume $\varphi^{\mathcal{M}}$, i.e., $(\exists u\psi(u))^{\mathcal{M}}$. Thus, for $u \in M$, $\psi^{\mathcal{M}}(u)$, and so by 1., $\psi(u)$. This entails $\exists u\psi(u)$, i.e., φ.
3. Let φ be $\forall u(\psi(u))$, ψ be Δ_0, and assume φ. Then, for all u, $\psi(u)$ and in particular, for all $u \in M$, $\psi(u)$. By 1. and $u \in M$, $\psi^{\mathcal{M}}(u)$. Therefore, $(\forall u \in M)\psi^{\mathcal{M}}(u)$, i.e., $\varphi^{\mathcal{M}}$.

□

The importance of Δ_0-absoluteness can be seen from the following:[47]

Lemma 3.26. The following notions are Δ_0, and hence absolute for transitive models: $x = \{u, v\}$, $x = (u, v)$, $x = \emptyset$, $x \subset y$, x is transitive, x is an ordinal, x is a limit ordinal, x is a natural number, $x = \omega$, $x = y \times z$, $x = y \setminus z$, $x = y \cap z$, $x = y \cup z$, R is a relation, f is a function, $x = \mathrm{dom}(R)$, $x = \mathrm{ran}(R)$, $x = f(y)$, $g = f \restriction y$. □

3.3.2 *Standard Models of Set Theory*

A standard model of ZFC is an uncountable, rank-initial segment of the cumulative hierarchy, that is closed under the set-of operations encoded by our current set theory (i.e., ZFC). The most important set-of operations are UNION, POWERSET, and REPLACEMENT. A level that satisfies these closure-properties will be labeled *inaccessible*.

Definition 3.27. A cardinal κ is *regular* iff there is no set $a < \kappa$ and function $f : a \to \kappa$ such that the image of f is unbounded in κ. A cardinal κ is *inaccessible*, if it is uncountable, regular, and $\lambda < \kappa$ implies $|\mathcal{P}(\lambda)| < \kappa$.

Regularity can be alternatively characterized as follows: κ is a regular cardinal if it cannot be expressed as the sum of fewer than κ-many smaller cardinals. So in order to get κ as the union of sets smaller than κ, we have to unite at least κ-many of them. The guiding idea is that we cannot reach κ from below by uniting less than κ-many smaller sets. Besides focusing only on uncountable cardinals, the notion of inaccessibility adds that we also cannot reach κ from below by applying the powerset-operation to sets smaller than κ.

Inaccessible ranks have some properties that will be important later on: if κ is inaccessible and $\gamma < \kappa$, then $|V_\gamma| < \kappa$ (which implies that $|V_\kappa| = \kappa$.) Moreover, for $a \in V_\kappa$, $|a| < \kappa$. Finally, if $a \subseteq V_\kappa$ such that $|a| < \kappa$, then $a \in V_\kappa$.[48]

The intuition behind an inaccessible cardinal is precisely that of being so large that it cannot be reached from below by application of the set-of operations of ZFC. Standard models of ZFC are inaccessible rank-initial segments of the cumulative hierarchy. If κ is the first inaccessible, V_κ is such a rank and hence, a model of ZFC and of it's second-order cousin ZFC_2.[49] Since I'm mostly interested in ZFC_2, I'll show it for the second-order case. The proof for the first-order case can easily be adopted. Note that the notion of logical consequence used for second-order theories here is always logical consequence in the sense of full semantics (see section 3.1.2):

47 For a proof, see [75, p. 164].

48 Proof sketch: if $\gamma = 0$, $|V_\gamma| = |V_0| < \kappa$. For $\gamma = \beta + 1$, suppose for IH that $|V_\beta| < \kappa$. Then $|V_\gamma| = |\mathcal{P}(V_\beta)| < \kappa$ since κ is inaccessible. Finally, if α is limit, suppose for IH that $|V_\gamma| < \kappa$ for all $\gamma < \alpha$. Then $|V_\gamma| < \kappa$ by regularity of κ. Now, if $a \in V_\kappa$, then $a \in V_\gamma$ for some $\gamma < \kappa$ since κ is a limit ordinal. Then, $a \subseteq V_\gamma$ by transitivity, and hence $|a| \leq |V_\gamma| < \kappa$. Finally, suppose that $a \subseteq V_\kappa$ such that $|a| < \kappa$. Then, each $b \in a$ enters the cumulative hierarchy below V_κ, and so $a \in V_\kappa$ since κ is inaccessible.

49 Strictly speaking, the model has the form $\langle V_\kappa, \in \cap (V_\kappa \times V_\kappa)\rangle$ for κ inaccessible, but I will simply use V_κ.

Theorem 3.28. If κ is inaccessible, then $V_\kappa \models \mathsf{ZFC}_2$.

Proof. The proof consists in verifying that all ZFC_2-axioms hold in V_κ:

EXTENSIONALITY Let $x, y \in V_\kappa$ be distinct. Then, there is a set $z \in x \wedge z \notin y$ (or *vice versa*), and since V_κ is transitive, $z \in V_\kappa$. Then,

$$V_\kappa \models \forall x \forall y (x \neq y \to \neg \forall u (u \in x \leftrightarrow u \in y))$$

which is equivalent to EXTENSIONALITY.

SEPARTION AXIOM Let $F \subseteq V_\kappa$ and $x \in V_\kappa$. Then $|x| < \kappa$. Moreover, by transitivity, $x \subseteq V_\kappa$, and since $x \cap F \subseteq x$, $x \cap F \subseteq V_\kappa$. Since $|x \cap F| \leq |x| < \kappa$, $x \cap F \in V_\kappa$.

PAIRING Let $x, y \in V_\kappa$. Then, $|x| < \kappa$ and $|y| < \kappa$, and hence there is an ordinal $\alpha < \kappa$ such that $x, y \in V_\alpha$. Then, $a = \{x, y\} \subset V_\alpha$, and so $a \in V_{\alpha+1}$. Since κ is inaccessible, $V_{\alpha+1} \subseteq V_\kappa$, and therefore $a \in V_\kappa$. Moreover, since the formula $x \in a \wedge y \in a$ is Δ_0, it holds in V_κ by Δ_0-absoluteness. Conversely, let $F(z) \equiv z = x \vee z = y$. Then, $\{z \in v \mid z = x \vee z = y\}$ exists by SEPARATION.

UNION Let $x \in V_\kappa$. Then, $|x| < \kappa$, and hence there is an ordinal $\alpha < \kappa$ such that $x \in V_\alpha$. Let $y \in \bigcup x$. Then, for some z, $y \in z \in x$. By transitivity, $z \in V_\alpha$ and also $y \in V_\alpha$. But then, $\bigcup x \in V_{\alpha+1}$, and hence, $\bigcup x \in V_\kappa$. By Δ_0, V_κ thinks that $\bigcup x$ contains all members of members of x.

POWERSET Let $x \in V_\kappa$. Then, $|x| < \kappa$, and hence there is an ordinal $\alpha < \kappa$ such that $x \in V_\alpha$. Let $y \in \mathcal{P}(x)$. Then $y \subseteq x$, and by transitivity of V_α, $y \subseteq V_\alpha$. So $y \in V_{\alpha+1}$ and hence, $\mathcal{P}(x) \in V_{\alpha+2}$ and therefore $\mathcal{P}(x) \in V_\kappa$. Moreover, since $\subseteq$ is Δ_0, V_κ thinks that $\mathcal{P}(x)$ contains all and only subsets of x.

INFINITY $\emptyset$ and ω are obviously in V_κ. Moreover, $\emptyset$ and successor are both Δ_0, and so V_κ thinks that ω contains $\emptyset$ and is closed under successor.

FOUNDATION Let $x \in V_\kappa$ such that $V_\kappa \models x \neq \emptyset$. Since $\emptyset$ is Δ_0, x is indeed nonempty by Δ_0-absoluteness. By metatheoretical FOUNDATION, there is some $y \in x$ such that $y \cap x = \emptyset$. Then, by transitivity of V_κ, $y \in V_\kappa$, and since $y \cap x = \emptyset$ is Δ_0, $V_\kappa \models y \cap x = \emptyset$.

REPLACEMENT AXIOM Let $G \subseteq V_\kappa \times V_\kappa$, that $x \in V_\kappa$, and that G is functional on x. Then $|x| < \kappa$, and if y is the image of x under G, then $|y| \leq |x| < \kappa$. Hence, $y \in V_\kappa$.

CHOICE Let $x \in V_\kappa$. Then, $x \times x$ as well as $\mathcal{P}(x \times x)$. By metatheoretical CHOICE, there is an $R \subset \mathcal{P}(x \times x)$ that is a well-ordering. Moreover, V_κ thinks that R is a linear ordering because the notion 'linear ordering' is Δ_0. Moreover, let $y \subseteq x$ so that V_κ thinks that y is non-empty. Then, as in the case of FOUNDATION, y is indeed non-empty, and by metatheoretical FOUNDATION, there is some $z \in y$ such that $z \cap x = \emptyset$. By Δ_0-absoluteness, $V_\kappa \models z \cap x = \emptyset$.

□

3.3.3 *Unintended Models of Set Theory*

As mentioned before, the first-order axiomatization of ZFC is not powerful enough to single out a unique class of models up to (rank-initial) isomorphism. One way to construct non-standard models of ZFC relies on the compactness theorem: let c be a fresh constant, and let s^n be the n-th iteration of the operation of von-Neumann successor ($s(x) = x \cup \{x\}$). Let $\mathsf{ZFC}^* := \mathsf{ZFC} \cup \{c \in \omega\} \cup \{s^n(\emptyset) \in c | n \in \omega\}$. Clearly, ZFC^* is finitely satisfiable and therefore, by the compactness theorem, satisfiable. But any model of ZFC which is also a model of ZFC^* will think that c is a finite von-Neumann ordinal that contains infinitely many elements.

Moreover, the Mostowski Collapse (theorem 3.22) has the aforementioned theorem 3.20, which implies the existence of non-standard models that truly deserve the name, as a corollary:

Theorem 3.18. Let $\mathcal{M} \models \mathsf{ZFC}$ be transitive. Then there exists a countable transitive model $\mathcal{N} \equiv \mathcal{M}$.

Proof. Let $\mathcal{M}$ be a transitive model of ZFC. By the Downward Löwenheim-Skolem Theorem (theorem 3.3), we obtain a countable elementary sub-structure $\mathcal{M}_0 \preceq \mathcal{M}$. Since $\mathcal{M}$ is transitive, $\in$ is well-founded on $\mathcal{M}$, and therefore also on $\mathcal{M}_0$. Since $\mathcal{M} \models \mathsf{ZFC}$ and $\mathcal{M}_0 \preceq \mathcal{M}$, $\mathcal{M}_0 \models$ Extensionality. We can therefore apply theorem 3.22 (Mostowski Collapse) to $\mathcal{M}_0$, and get an isomorphic (and hence, countable) transitive structure $\mathcal{N}$. Taken together, this yields $\mathcal{N} \cong \mathcal{M}_0 \preceq \mathcal{M}$, and hence $\mathcal{N} \equiv \mathcal{M}$. □

It is important to note that $\mathcal{N}$ is a *countable transitive* model of ZFC. However, by ZFC, there are uncountable sets, so $\mathcal{N} \models$ 'there are uncountable sets'. Since $\mathcal{N}$ is transitive, the uncountably-many elements must be contained in $\mathcal{N}$.

3.3.4 *Quasi-Categoricity of Second-Order Set Theory*

By theorem 3.28, if κ is inaccessible, $V_\kappa \models \mathsf{ZFC}_2$ (and also ZFC). In the second-order case, the converse holds as well: if $V_\kappa \models \mathsf{ZFC}_2$, then κ is inaccessible. This is crucial for the quasi-categoricity theorem of ZFC_2. After having established this crucial result, I will go on and show that any ZFC_2-model is isomorphic to a strongly inaccessible rank of V.

Theorem 3.27. If $V_\kappa \models \mathsf{ZFC}_2$, then κ is an inaccessible cardinal.

Proof. The proof has four steps:

1. κ is a cardinal: Suppose otherwise. Then, there is a bijection $f : \beta \to \kappa$ for some $\beta < \kappa$. Since $\beta \in V_\kappa$, Replacement implies that $\kappa \in V_\kappa$ as well, contradicting the fact that $\mathrm{rank}(\kappa) = \kappa$.

2. $\kappa > \omega$: Let $\alpha \in V_\kappa$ be such that V_κ thinks that α is an infinite ordinal. (Such an α exists because $V_\kappa \models$ Infinity.) Since 'being an ordinal' is Δ_0 (and V_κ is transitive), the notion is absolute, so α is an ordinal. Similarly, the empty set and the successor operation are absolute, so α is infinite. Since $\alpha \in V_\kappa$, κ is infinite, too.

3. κ is a regular cardinal: Let $f : \beta \to \kappa$, for a $\beta < \kappa$ and let $a = f(\beta)$. Then, $a \in V_\kappa$, by REPLACEMENT within V_κ. Since κ is a limit ordinal, there is a $\gamma < \kappa$ such that $a \in V_\gamma$, i.e., $a \subseteq \gamma$, and f's range is bounded. Then, κ is regular.

4. $\lambda < \kappa \Rightarrow |\mathcal{P}(\lambda)| < \kappa$:
   a) The first step is to show that for any $a \in V_\kappa$, $\mathcal{P}^{V_\kappa}(a) = \mathcal{P}(a)$: $\mathcal{P}^{V_\kappa}(a) \subseteq \mathcal{P}(a)$ holds by transitivity and the absoluteness of subset, as in theorem 3.28. Next, I show that $\mathcal{P}(a) \subseteq \mathcal{P}^{V_\kappa}(a)$. Suppose $b \subseteq a$. So, by transitivity $b \subseteq a \subseteq V_\kappa$. By SEPARATION and the absoluteness of intersection, $b = (a \cap b) \in V_\kappa$. Hence, indeed $\mathcal{P}(a) \subseteq \mathcal{P}^{V_\kappa}(a)$ for elements $a \in V_\kappa$.
   b) Since V_κ satisfies CHOICE, it holds that in V_κ, $\mathcal{P}^{V_\kappa}(\lambda) = \mathcal{P}(\lambda)$ can be mapped bijectively to some ordinal. Since this notion is absolute, $\mathcal{P}(\lambda)$ is indeed in a bijection with some ordinal in V_κ, i.e., some ordinal smaller than κ. Therefore, $|\mathcal{P}(\lambda)| < \kappa$.

□

Before I can show that ZFC_2-models are isomorphic to ranks of the cumulative hierarchy, I need a technical lemma that is mostly due to FOUNDATION, and that I state without proof:[50]

Lemma 3.28. ZFC_2 proves: if $F \neq \emptyset$, then $\exists v(Fv \wedge (\forall x \in v)\neg Fx)$. □

This allows me to prove the following theorem, which is important for the proof of the quasi-categoricity theorem:

Theorem 3.29. $\mathcal{M} \models \mathsf{ZFC}_2$ if and only if $\mathcal{M} \cong V_\kappa$ for some inaccessible κ.

Proof. $\Rightarrow$ I show that if $\mathcal{M} \models \mathsf{ZFC}_2$, then $\mathcal{M} \cong V_\alpha$ for some ordinal α. This is mainly due to theorem 3.22 (Mostowski Collapse). It then follows by theorem 3.27 that α is an inaccessible cardinal.

First, I need to check that I can apply the Mostowski Collapse. Since $\mathcal{M} \models \mathsf{ZFC}_2$, it is extensional. To see that $\mathcal{M}$ is also well-founded, assume that there is an infinite descending $\in$-chain in $\mathcal{M}$, i.e., there is a sequence of elements a_n for $n \in \mathbb{N}$, so that $\mathcal{M} \models a_{n+1} \in a_n$. Now, let $F = \{a_n \in M \mid n \geq 0\}$. By lemma 3.28, there must be some $x \in a_n \in F$ for some $i \geq 0$, s.t. $x \cap F = \emptyset$. But this contradicts the assumption.

Since $\mathcal{M}$ is extensional and well-founded, I can apply the Mostowski-Collapse. Then, there is a transitive structure $\mathcal{N}$ s.t. $\mathcal{M} \cong \mathcal{N}$. From now on, I work in $\mathcal{N}$ to show that $\mathcal{N} = V_\alpha$ for some ordinal α.

The first step is to show that, $\mathcal{P}^{\mathcal{N}}(a) = \mathcal{P}(a)$ for any $a \in N$, which is mostly analogous to step 4.a) in the proof of theorem 3.27: $\mathcal{P}^{\mathcal{N}}(a) \subseteq \mathcal{P}(a)$ holds by transitivity and absoluteness of subset. For $\mathcal{P}(a) \subseteq \mathcal{P}^{\mathcal{N}}(a)$, let $b \subseteq a$. By SEPARATION, $b \cap a = b$ exists in $\mathcal{N}$. So $\mathcal{P}^{\mathcal{N}}(a) = \mathcal{P}(a)$. Then $V_\gamma^{\mathcal{N}} = V_\gamma$ follows by induction on γ in $\mathcal{N}$.

50 A proof is given in Button and Walsh [18, p. 191].

Now let α be the least ordinal not in $\mathcal{N}$. I show that $\mathcal{N} = V_\alpha$. Suppose that $a \in \mathcal{N}$. Since $\mathcal{N} \models \mathsf{ZFC}_2$, it follows that $a \in V_\gamma^{\mathcal{N}} = V_\gamma$ for some ordinal $\gamma < \alpha$. Therefore, $a \in V_\alpha$. Now let $a \in V_\alpha$. Again, since $\mathcal{N} \models \mathsf{ZFC}_2$, α is a limit. Therefore, $a \in V_\gamma$ for some $\gamma < \alpha$, and hence $a \in V_\gamma = V_\gamma^{\mathcal{N}} \subseteq \mathcal{N}$.

$\Leftarrow$ Let $\mathcal{M} \cong V_\kappa$ for some inaccessible κ. Then, by theorem 3.28, $\mathcal{M} \models \mathsf{ZFC}_2$. □

Now we can finally proof Zermelo's [173] famous *Quasi-Categoricity Theorem*:

Theorem 3.19. Let $\mathcal{M} \models \mathsf{ZFC}_2$ and $\mathcal{N} \models \mathsf{ZFC}_2$. Then exactly one of the following obtains:

(i) $\mathcal{M} \cong \mathcal{N}$

(ii) $\mathcal{M} \cong V_\alpha^{\mathcal{N}}$ for some α which $\mathcal{N}$ thinks is an inaccessible cardinal.

(iii) $\mathcal{N} \cong V_\alpha^{\mathcal{M}}$ for some α which $\mathcal{M}$ thinks is an inaccessible cardinal.

Proof. By theorem 3.29, there are inaccessible cardinals κ and λ so that $\mathcal{M} \cong V_\kappa$ and $\mathcal{N} \cong V_\lambda$. There are three cases:

1. $\kappa = \lambda$. Then $V_\kappa = V_\lambda$, which implies $V_\kappa \cong V_\lambda$, and so we get $\mathcal{M} \cong V_\kappa \cong V_\lambda \cong \mathcal{N}$. Since $\cong$ is transitive, (i) follows.

2. $\kappa < \lambda$. This implies that $\kappa \in V_\lambda$. Since 'being a ordinal' is absolute, κ is an ordinal according to V_λ. Since κ is regular, this implies that there cannot be a $\gamma < \kappa$ and a function $f : \gamma \to \kappa$ that has an unbounded range. Since this property ('being a function $f : \gamma \to \kappa$ with unbounded range') is also absolute, there cannot be such a function in V_λ, which implies that V_λ thinks that κ is regular. Finally, the inaccessibility of κ implies that $|\mathcal{P}(\gamma)| < \kappa$ for any $\gamma < \kappa$. We can therefore apply theorem 3.28, and get that $|\mathcal{P}^{V_\lambda}(\gamma)| = |\mathcal{P}(\gamma)| < \kappa$. Finally, theorem 3.28 implies that V_λ thinks that the powerset of γ has cardinality $< \kappa$. So V_λ thinks that κ is inaccessible. So we get that $\mathcal{M} \cong V_\kappa \cong V_\alpha^{\mathcal{N}}$, and therefore (ii) obtains.

3. As in 2., establishing (iii).

□

3.4 THE RAYO-UZQUIANO APPROACH

In this section, I'll introduce the *Rayo-Uzquiano Approach* (RU approach). The RU approach offers an alternative to standard model theory by using higher-order logic. Unlike standard model theory, which is developed in the language of set theory (or a set-theoretically enriched version of (first-order) mathematical English), Rayo and Uzquiano [135] develop their approach strictly within the framework of higher-order logic.

The main advantage of the RU approach is its ability to overcome the restriction to set-sized domains.[51] This restriction are problematic for the absolutist.

51 What counts as "set-sized" depends on the set-theoretic background assumptions: for instance, in ZFC− Infinity, every set is finite; in ZFC, every set is smaller than the first inaccessible cardinal, and so on.

As we saw in chapter 2, sets cannot be absolutely general—at least within the framework of standard iterative set theory and classical logic. As a result, no model-theoretic domain can achieve absolute generality. However, by combining higher-order logic with a non-standard interpretation of higher-order expressions, the RU approach bypasses these set-theoretic size restrictions. However, much of this promise hinges on the philosophical interpretation of the higher-order expressions.

As we saw in section 2.2.3, one option is to adopt a roughly Fregean approach based on concepts to interpret higher-order expressions. In this approach, a universal property (such as self-identity) can serve as the domain for unrestricted quantification. Alternatively, we can turn to the *plural interpretation* developed by George Boolos in the 1980s.[52] According to this interpretation, the *plurality of absolutely everything* can be taken as the domain for unrestricted quantification. This is merely a suggestion for how to interpret higher-order expressions in a way that is compatible with generality absolutism. However, in this section, I will focus on the technical presentation of the material and largely postpone the philosophical discussion of these alternative interpretations to the following chapters.[53] As a heuristic principle, we will assume for now that model-theoretic notions, once translated into the higher-order framework, are compatible with generality absolutism. Whether this assumption is correct, and what philosophical implications it may have, will be among the central topics of chapter 6.

The core idea of the RU approach is as follows: fix an object language. This language is typically equipped with the ability to name objects and express relational claims. To facilitate this, the language must include resources for denoting objects (such as constants and variables) and relations. The interpretation of these linguistic expressions consists of mappings between the expressions and the some domain. Specifically, the interpretation of a constant corresponds to a simple term-object relation, while the interpretation of a relation constant is a more complex relation-objects relation, where the interpretation itself is a higher-type relation than the one it interprets.

Higher-order logic serves as the ideal tool for modeling these interpretation relations because, unlike first-order logic, it has the resources to quantify over relations. Type-0 terms denote objects, and thus their interpretations are type-1 relations between type-0 terms and objects in the domain. Type-1 terms denote predicates or relations between type-0 terms, and their interpretations are type-2 relations between type-1 terms and tuples of objects. More generally, interpretations for n^{th}-order object languages are relations between n^{th}-order expressions and the objects in the domain. This is formalized in an $(n+1)^{\text{th}}$-order metalanguage.

In my view, the RU approach can best be understood as an alternative to standard model theory, wherein we employ a different formal tool—higher-order logic—but retain the same fundamental concepts and principles (such as the structure of models, that satisfaction is a relation between a language and the world, and so on). Consequently, the primary task of this section is to

52 See Boolos [16, 12] and section 3.1.3.

53 See in particular section 6.1.

start with model-theoretic concepts and translate them into higher-order logic. According to the heuristic principle outlined above, this translation makes the higher-order analogues of model-theoretic notions compatible with generality absolutism *for the moment*.

The RU approach was first developed for the first- and second-order languages of set theory in Rayo and Uzquiano [135], and later generalized to arbitrary first- and second-order languages in Rayo and Williamson [136]. I will adopt the terminology and notation used in Rossi [144]. Not only will I introduce the RU approach, but I will also extend it in two significant directions: first, I will incorporate higher-order object languages of arbitrary (finite) order.[54] Second, I will expand the approach by using both standard Tarski satisfaction and, following Rossi [144], Kleene satisfaction. Additionally, in contrast to the work of the aforementioned authors, I will develop the RU approach within a type-theoretic framework.

To begin, I will describe the languages used as object- and metalanguages. Then, I will define the higher-order counterparts of the following model-theoretic notions: 'model' ($\mathbb{M}$), 'domain' ($\mathbb{D}$), 'variation of a variable assignment' ($\mathbb{V}$), and 'satisfaction' (Sat).

3.4.1 *The Languages*

As pointed out before (section 3.1.2), I use a hierarchy of higher-order languages (simple type theory), which allows for cross-type predication (i.e., the type theory is cumulative), and adopt the respective notational conventions. So type-0 terms stand for objects, type-1 terms can be applied to objects, type-2 terms can be applied to type-0 *and* type-1 terms, and so on. This leads to the following definition:[55]

Definition 3.31. Let $\mathfrak{L} := \mathcal{L}^1, \mathcal{L}^2, ..., \mathcal{L}^n, ...$ be a sequence of countable higher-order languages such that each $\mathcal{L}^i \in \mathfrak{L}$ satisfies the following requirements:

1. $\mathcal{L}^i$ contains $\in$ but has no function symbols.
2. $\mathcal{L}^i$ and $\mathcal{L}^{i+1}$ have the same (variable and constant) terms up to type $(i-1)$, but $\mathcal{L}^{i+1}$ additionally has (variable and constant) terms of type-i.
3. For all $i \geq 2$, $\mathcal{L}^i$ contains the symbols $\langle \cdot \rangle$ and $\ulcorner \cdot \urcorner$.

The first requirement ensures that every language contains the symbol $\in$, which allows mathematical concepts to be expressed in their set-theoretic representation. Additionally, the exclusion of function symbols is purely for simplicity, and it imposes no real cost, since functions can always be defined as

54 Although there seems to be no principled reason to limit the approach to finite-order languages, I will focus primarily on finite-order languages here, as they suffice for the purposes of this discussion. For a treatment of a transfinite hierarchy of metalanguages and their connection to generality absolutism, see Linnebo and Rayo [93].

55 At this point, I only give the rare backbone of the semantic theory without any topic-specific aspects. This means that in the presentation of the languages, I do not discuss specific non-logical expressions that will be relevant later on such as a truth predicate (chapter 4) or the membership relation (chapter 5).

relations that are univocal to the right. The second requirement ensures that, as we ascend the hierarchy, languages differ only in that new terms are introduced. Specifically, we have $\mathcal{L}^i \subseteq \mathcal{L}^{i+1}$, and once a term is added, it is never removed. The third requirement ensures that each language from $\mathcal{L}^2$ onward contains primitive symbols designed to internalize the notions of *denotation* and *being in the quantifier domain*. I will explain how these expressions are used in the following section.

3.4.2 *The RU model*

I will now define the notion of a 'model for an n^{th}-order object language' by means of a predicate, $\mathbb{M}(X^n)$, in an $(n+1)^{\text{th}}$-order metalanguage. This predicate applies to a free unary type-n variable X^n (recall that an $(n+1)^{\text{th}}$-order language has variables of type at most n). Capital letters are used in definitions solely to enhance readability.

In classical model theory, first-order models serve the following four purposes: (1) providing a domain of individuals, (2) interpreting individual constants and relation symbols, (3) denoting terms, and (4) reasoning about satisfaction. A model for a first-order language $\mathcal{M} = \langle M, I \rangle$ accomplishes this in the following way: the domain is represented by a non-empty set M, and the interpretation is represented by the interpretation function I. Terms, which may be individual constants or variables (assuming languages without function symbols), are denoted by the interpretation function (in conjunction with a variable assignment).

Let us begin with a language that only has type-0 terms, serving as a toy example. (This example is purely heuristic, as such a language is relatively uninteresting since it can only name objects.) A RU-model for a type-0 object language is defined in a metalanguage with type-1 terms (i.e., a second-order language). A semi-formal characterization looks as follows: for each unary type-1 variable X^1, the formula "X^1 is an RU-model for a type-0 object language," denoted $\mathbb{M}(X^1)$, is defined as:

$$\begin{aligned}\mathbb{M}(X^1) :\leftrightarrow\ & X^1 \text{ is non-empty} \wedge\ X^1 \text{ contains only interpretational objects} \wedge \\ & \forall x^0 \big[\text{if } X^1(x^0) \text{ then } x^0 \text{ is in the quantifier domain}\big] \wedge \\ & \forall z^0 \big[\mathsf{Con}^0(z^0) \vee \mathsf{Var}^0(z^0), X^1 \text{ contains an interpretation of } z^0\big]\end{aligned}$$

The first conjunct is simply the higher-order equivalent of the non-emptiness requirement in standard model theory. Let me now explain the third and fourth conjuncts, and I will return to the second one shortly. By 'X^1 contains an interpretation of z^0', I mean that there exists an object in X^1 that interprets z^0. The notions of 'being in the quantifier domain' and 'y^0 interprets z^0' (or 'z^0 denotes y^0'), where y^0 is an object in the domain and z^0 is an expression in the language, are formalized in the spirit of Rayo and Uzquiano [135] by using the primitive expressions $\langle \cdot \rangle$ and $\ulcorner \cdot \urcorner$ from definition 3.31 as follows:

- $\langle \ulcorner \forall \urcorner, y^1 \rangle$ means 'y^0 is in the quantifier domain', or 'y^0 can be quantified over',

- $\langle z^0, y^0 \rangle$ means 'z^0 denotes y^0'.

Note that, since we are working with a type-0 language, the union of constant and variable type-0 terms ($\forall z^0 \big[\mathsf{Con}^0(z^0) \vee \mathsf{Var}^0(z^0) \ldots$) completely exhausts the (non-logical) resources of the object language. Therefore, X^1 contains tuples that encode the information that some objects are in the quantifier domain, and for each expression in the language, X^1 contains a tuple that codes its interpretation. I will refer to tuples that code interpretations as *interpretational objects*. Thus, X^1 combines the domain, the interpretation function, and the variable assignment (requirements 1 – 3).[56]

The second conjunct, which states that 'everything X^1 contains is an interpretational object', expresses that models apply only to interpretational objects. I refer to this as the *interpretational objects requirement*.[57] Translating the model predicate for languages with only type-0 terms into the higher-order framework yields the following:

$$\begin{aligned}
&\mathbb{M}(X^1) :\leftrightarrow \exists x^0 X^1(\langle \ulcorner \forall \urcorner, x^0 \rangle) \wedge \\
&\quad \forall x^0 \Big(X^1(x^0) \rightarrow \big(\exists y^0 (x^0 = \langle \ulcorner \forall \urcorner, y^0 \rangle) \vee \exists z^0 \exists y^0 (x^0 = \langle z^0, y^0 \rangle) \big) \Big) \wedge \\
&\quad \forall z^0 \Big(\mathsf{Con}^0(z^0) \vee \mathsf{Var}^0(z^0) \rightarrow \exists! y^0 (X^1(\langle z^0, y^0 \rangle) \wedge X^1(\langle \ulcorner \forall \urcorner, y^0 \rangle)) \Big)
\end{aligned}$$

The first conjunct ensures non-emptiness. The second conjunct expresses the *interpretational objects requirement*, stating that every element in X^1 either encodes that something is in the quantifier domain ($\exists y^0 (x^0 = \langle \ulcorner \forall \urcorner, y^0 \rangle)$), or encodes an interpretation of a (constant or variable) type-0 term ($\exists z^0 \exists y^0 (x^0 = \langle z^0, y^0 \rangle)$). The third conjunct ensures that every (constant or variable) type-0 term in the language is interpreted by some object y^0, and that y^0 can be quantified over.

Let us now consider a more interesting second-order object language. The interpretation is formalized as a predicate in a third-order metalanguage, which contains constant and variable terms up to type 2. The model predicate applies to a free unary type-2 variable X^2. In classical model theory, an interpretation function for a second-order language $\mathcal{L}^2$ encompasses all first-order expressions, in addition to the second-order variables. If $\mathcal{L}^2$ is a second-order language and $\mathcal{L}^1$ is its first-order fragment, then we have $\mathcal{L}^1 \subseteq \mathcal{L}^2$. Moreover, if I^2 is a $\mathcal{L}^2$-interpretation, then the restriction of I^2 to $\mathcal{L}^1$, denoted $I^2 \restriction_{\mathcal{L}^1}$, is a $\mathcal{L}^1$-interpretation such that $I^2 \restriction_{\mathcal{L}^1} \subseteq I^2$.

Note that every second-order interpretation I^2 has a unique restriction to its first-order fragment, and every first-order interpretation function I^1 can be extended (in multiple ways) to an interpretation function for an extension of the language $\mathcal{L}^2 \supseteq \mathcal{L}^1$. To encode the subset relation, I use the following abbreviation:

$$X^1 \sqsubseteq X^2 :\leftrightarrow \forall x^0 (X^1(x^0) \rightarrow X^2(x^0)).$$

Here, $\sqsubseteq$ is the higher-order analogue of the set-theoretic subset relation $\subseteq$. To encode that every $\mathcal{L}^2$-interpretation I^2 has a corresponding $\mathcal{L}^1$-restriction $I^2 \restriction_{\mathcal{L}^1}$,

56 We will address requirement 4, satisfaction, in section 3.4.4.

57 This requirement will be crucial for the proof of lemma 3.36.

which stands in the subset relation to I^2, we use a type-1 variable X^1 that is in the $\sqsubseteq$-relation with the model variable X^2. Note that the definition of $\sqsubseteq$ could not be given in strict type theory, since $X^2(x^0)$ would not be grammatically correct.[58]

The choice to formulate RU-interpretations of second-order languages (an of nth-order lanaguages more generally) in a *cumulative* type theory is not strictly necessary since it is possible to rephrase all relevant definitions within strict type theory. On such an approach, however, the relationship between a second-order model and its first-order reduct takes a different form. In particular, the first-order reduct would not be a *part* of the second-order model in the sense that the second-order model X^2 could directly apply to the interpretational objects of the first-order model (i.e., to type-o objects as $X^2(x^0)$ which is permitted in the cumulative setting and used in the definition of $\sqsubseteq$). Instead, the second-order model would only contain the first-order model *as a complete object*, i.e., via $X^2(X^1)$. I adopt the cumulative approach because its closer resemblance to the standard model-theoretic treatment of second-order logic, where interpretation functions are typically nested in the sense that they contain interpretations for expressions of lower type.

The first part of the definition of an RU-model for a second-order object language is identical to the definition given above, where X^1 represents the restriction of the interpretation to the fragment dealing only with type-0 terms. The part of the interpretation that deals with type-1 terms is encoded in X^2. In semi-formal notation, the definition of a model for a second-order object language, denoted $\mathbb{M}(X^2)$, is as follows:

$$\begin{aligned}\mathbb{M}(X^2) :\leftrightarrow\ & \exists X^1 \left(X^1 \sqsubseteq X^2\right) \wedge \\ & \Big[X^1 \text{ is non-empty} \wedge \text{ everything in } X^1 \text{ is an interpretational object} \wedge \\ & \quad X^1 \text{ contains interpretations for all type-0 terms}\Big] \wedge \\ & \big[X^2 \text{ is non-empty} \wedge \text{ everything in } X^2 \text{ is an interpretational object} \wedge \\ & \quad X^2 \text{ contains all interpretations for all type-1 terms}\big]\end{aligned}$$

Where z^1 is an n-ary type-1 term, and $y^0_1, \ldots, y^0_n$ are type-0 terms, the fact that $y^0_1, \ldots, y^0_n$ are z^1-related is formalized by the expression $\langle z^1, \langle y^0_1, \ldots, y^0_n\rangle\rangle$, which we read as '$\langle y^0_1, \ldots, y^0_n\rangle$ is in the interpretation of z^1'. Here is the formal definition of the notion 'model for a second-order object language', an explanatory remark follows afterwards:

Definition 3.32. For every unary type-2 variable X^2, the formula "X^2 is an RU-Model for a type-1 object language", in symbols $\mathbb{M}(X^2)$, is defined as:

$$\begin{aligned}\mathbb{M}(X^2) :\leftrightarrow\ & \exists X^1(X^1 \sqsubseteq X^2) \wedge \Big[\exists x^0 X^1(\langle {}^\ulcorner\forall^\urcorner, x^0\rangle) \wedge \\ & \forall x^0\Big(X^1(x^0) \rightarrow \big(\exists y^0(x^0 = \langle {}^\ulcorner\forall^\urcorner, y^0\rangle) \vee \exists z^0 \exists y^0(x^0 = \langle z^0, y^0\rangle)\big)\Big) \wedge \\ & \forall z^0\Big(\mathsf{Con}^0(z^0) \vee \mathsf{Var}^0(z^0) \rightarrow \exists! y^0(X^1(\langle z^0, y^0\rangle) \wedge X^1(\langle {}^\ulcorner\forall^\urcorner, y^0\rangle))\Big)\Big] \wedge\end{aligned}$$

58 Recall that in strict type theory, $x^n(x^m)$ is grammatical only if $n = m + 1$ (see section 3.1.2).

$$\Big[\exists x^1 X^2(\langle\ulcorner\forall\urcorner, x^1\rangle) \wedge \forall x^0(X^2(x^0) \to X^1(x^0)) \wedge$$
$$\forall x^1\Big(X^2(x^1) \to \big(\exists y^1(x^1 = \langle\ulcorner\forall\urcorner, y^1\rangle) \vee \exists y^1 \exists y_1^0, ..., \exists y_m^0(x^1 = \langle y^1, \langle y_1^0, ..., y_m^0\rangle\rangle)\big)\Big) \wedge$$
$$\forall z^1\Big(\mathsf{Con}^1(z^1) \vee \mathsf{Var}^1(z^1) \to \exists y_1^0, ..., \exists y_m^0(X^2(\langle z^1, \langle y_1^0, ..., y_m^0\rangle\rangle)) \wedge$$
$$\forall y_1^0, ..., y_m^0(X^2(\langle z^1, \langle y_1^0, ..., y_m^0\rangle\rangle) \to X^1(\langle\ulcorner\forall\urcorner, y_1^0\rangle) \wedge ... \wedge X^1(\langle\ulcorner\forall\urcorner, y_m^0\rangle))\Big)\Big]$$

Remark 3.33.

- The first large square bracket contains a copy of $\mathbb{M}(X^1)$, where X^1 is used to interpret type-0 terms and is in the $\sqsubseteq$-relation with X^2.
- In the second large square bracket, the type-2 variable X^2 is used to interpret both constant and variable type-1 terms. The first conjunct repeats the *non-emptiness requirement*. The next two conjuncts are necessary for the *interpretational objects requirement*. The first of these asserts that there are no non-interpretational type-0 objects in X^2, while the second specifies the type-1 interpretational objects that are added to X^2.
- The final two lines express that every type-1 term must have an interpretation in X^2. Since relations apply to objects, we use $X^1(\langle\ulcorner\forall\urcorner, y_1^0\rangle) \wedge \cdots \wedge X^1(\langle\ulcorner\forall\urcorner, y_m^0\rangle)$, where X^1 replaces X^2, to express that the objects are part of the first-order domain.
- A model for a classical first-order language (with constant and variable type-0 terms but only constant type-1 terms) would essentially resemble the definition above, except that the set of variable type-1 terms would be empty.

Note that, even though I explicitly required in definition 3.31 that every language $\mathcal{L}^n \in \mathfrak{L}$ contains $\in$ as a primitive expression, the interpretation of $\in$ is not explicitly mentioned in definition 3.32. This omission is intentional: $\in$ is a constant type-1 term, and its interpretation is implicitly covered by the corresponding clause. This approach generalizes the original one of Rayo and Uzquiano [135], who considered only the language of (first- and second-order) set theory as their object language. In the current work, however, I consider arbitrary n^{th}-order object languages.

The definition above already suggests how to generalize the model predicate to higher-order object languages: whenever we have an RU-model for an $(n-1)^{\text{th}}$-order object language as a definition in an n^{th}-order metalanguage, and we wish to define an RU-model for an n^{th}-order object language, we must transition to an $(n+1)^{\text{th}}$-order metalanguage and add a conjunct in which a variable type-$(n+1)$ term is used to interpret the type-n expressions.

However, when transitioning from models for second-order to third-order object languages, there is another important aspect to consider. Recall that I employ a cumulative, rather than strict, type theory; that is, predication $a^k(b^l)$ is well-formed if and only if $l < k$. The distinction between cumulative and strict type theory is irrelevant in the case of second-order object languages, as

there are only type-0 terms to which type-1 terms can be applied. However, once we move to a third-order object language, we can encounter cross-type predications such as $a^2(b^0, c^1)$.

I now turn definition 3.32 into a definition of a 'model for a third-order object language'. First, I need to replace the first conjunct with the following:

$$\exists X^1 \exists X^2 (X^1 \sqsubseteq X^2 \wedge X^2 \sqsubseteq X^3)$$

where X^2 contains all interpretations encoded in X^1 and X^3 contains all interpretations coded in X^2 and X^1 (note that $\sqsubseteq$ is transitive). This generalizes the earlier observation about model-theoretic interpretation functions. Whenever we have an n^{th}-order language $\mathcal{L}^n$, and we restrict the corresponding interpretation function I^n to the i^{th}-order fragment (for $i < n$), the resulting function $I^n \restriction_{\mathcal{L}^i}$ is a $\mathcal{L}^i$-interpretation and a subset of I^n.

Second, I need to add a conjunct interpreting type-2 terms. Let $k, l \in \{0, 1\}$, and define the conjunct as follows:

$$\Big[\exists x^2 X^3(\langle \ulcorner \forall \urcorner, x^2 \rangle) \wedge \forall x^1 (X^3(x^1) \rightarrow X^2(x^1)) \wedge$$
$$\forall x^2 \Big(X^3(x^2) \rightarrow \big(\exists y^2 (x^2 = \langle \ulcorner \forall \urcorner, y^2 \rangle) \vee \exists y^2 \exists y^k_1, ..., \exists y^l_m (x^2 = \langle y^2, \langle y^k_1, ..., y^l_m \rangle \rangle)\big)\Big) \wedge$$
$$\forall z^2 \Big(\mathsf{Con}^2(z^2) \vee \mathsf{Var}^2(z^2) \rightarrow \exists y^k_1, ..., \exists y^l_m (x^2 = \langle z^2, \langle y^k_1, ..., y^l_m \rangle \rangle) \wedge$$
$$\forall y^k_1, ..., y^l_m (X^3(\langle z^2, \langle y^k_1, ..., y^l_m \rangle \rangle) \rightarrow X^{k+1}(\langle \ulcorner \forall \urcorner, y^k_1 \rangle) \wedge \cdots \wedge X^{l+1}(\langle \ulcorner \forall \urcorner, y^l_m \rangle))\Big)\Big]$$

The main change in the above conjunct is the use of the variables $k, l \in \{0, 1\}$ to indicate the type of objects to which a type-2 term applies.

Let me now provide the full definition of the notion of a 'model for an n^{th}-order object language', as defined in an $(n+1)^{\text{th}}$-order metalanguage. As expected, the definition will consist of n conjuncts, each of which addresses the interpretation of a particular type of expression. Cross-type predication concerns each type > 1, so each conjunct for expressions of type > 1 uses variables for the corresponding types. Since the full definition is lengthy—and can only be given for a fixed n—I will provide an abbreviated version. The resulting definition is as follows:

Definition 3.34. Let $i, j \in \{0, ..., n-1\}$. For every unary type-$(n+1)$ variable X^{n+1}, we define the formula "X^{n+1} is a RU-Model for a type-n object language", in symbols $\mathbb{M}(X^{n+1})$, as:

$$\mathbb{M}(X^{n+1}) :\leftrightarrow \exists X^1 \exists X^2 \exists X^3 ... \exists X^n (X^1 \sqsubseteq X^2 \wedge X^2 \sqsubseteq X^3 \wedge ... \wedge X^n \sqsubseteq X^{n+1}) \wedge$$
$$\Big[\exists x^0 X^1(\langle \ulcorner \forall \urcorner, x^0 \rangle) \wedge \forall x^0 \Big(X^1(x^0) \rightarrow \big(\exists y^0 (x^0 = \langle \ulcorner \forall \urcorner, y^0 \rangle) \vee \exists z^0 \exists y^0 (x^0 = \langle z^0, y^0 \rangle)\big)\Big) \wedge$$
$$\forall z^0 \Big(\mathsf{Con}^0(z^0) \vee \mathsf{Var}^0(z^0) \rightarrow \exists! y^0 (X^1(\langle z^0, y^0 \rangle) \wedge X^1(\langle \ulcorner \forall \urcorner, y^0 \rangle))\Big)\Big] \wedge$$
$$\Big[\exists x^1 X^2(\langle \ulcorner \forall \urcorner, x^1 \rangle) \wedge \forall x^0 (X^2(x^0) \rightarrow X^1(x^0)) \wedge$$
$$\forall x^1 \Big(X^2(x^1) \rightarrow \big(\exists y^1 (x^1 = \langle \ulcorner \forall \urcorner, y^1 \rangle) \vee \exists y^1 \exists y^0_1, ..., \exists y^0_m (x^1 = \langle y^1, \langle y^0_1, ..., y^0_m \rangle \rangle)\big)\Big) \wedge$$
$$\forall z^1 \Big(\mathsf{Con}^1(z^1) \vee \mathsf{Var}^1(z^1) \rightarrow \exists y^0_1, ..., \exists y^0_m (X^2(\langle z^1, \langle y^0_1, ..., y^0_m \rangle \rangle)) \wedge$$

$$\forall y_1^0,...,y_m^0(X^2(\langle z^1,\langle y_1^0,...,y_m^0\rangle\rangle)\rightarrow X^1(\langle\ulcorner\forall\urcorner,y_1^0\rangle)\wedge...\wedge X^1(\langle\ulcorner\forall\urcorner,y_m^0\rangle))\Big)\Big]\wedge...\wedge$$

$$\Big[\exists x^n X^{n+1}(\langle\ulcorner\forall\urcorner,x^n\rangle)\wedge\forall x^{n-1}(X^{n+1}(x^{n-1})\rightarrow X^n(x^{n-1}))\wedge$$

$$\forall x^n\Big(X^{n+1}(x^n)\rightarrow(\exists y^n(x^n=\langle\ulcorner\forall\urcorner,y^n\rangle)\vee\exists y^n\exists y_1^i,...,\exists y_m^j(x^n=\langle y^n,\langle y_1^i,...,y_m^j\rangle\rangle))\Big)\wedge$$

$$\forall z^n\Big(\mathsf{Con}^n(z^n)\vee\mathsf{Var}^n(z^n)\rightarrow\exists y_1^i,...,\exists y_m^j(x^n=\langle z^n,\langle y_1^i,...,y_m^j\rangle\rangle)\wedge$$

$$\forall y_1^i,...,y_m^j(X^{n+1}(\langle z^n,\langle y_1^i,...,y_m^j\rangle\rangle)\rightarrow X^{i+1}(\langle\ulcorner\forall\urcorner,y_1^i\rangle)\wedge...\wedge X^{j+1}(\langle\ulcorner\forall\urcorner,y_m^j\rangle))\Big)\Big]$$

Remark 3.35 (Polyadic Quantification)**.** Although the metatheory used to develop the RU approach only uses monadic plural quantification, this does not restrict the object theory to monadic higher-order quantification. The framework includes a primitive pairing operation of the form $\langle z^0, y^0\rangle$, which is used to express that z^0 denotes y^0. These pairing expressions can be freely nested and extended, enabling us to represent tuples of arbitrary length. In particular, interpretations for n-adic higher-order variables are encoded by nested pairs of the form $\langle y^s,\langle y_1^k,\ldots,y_n^l\rangle\rangle$ $(k,l<s)$, where y^s functions as an n-ary relation. As such, the system can mimic polyadic quantification in the object language despite working with a monadic plural metatheory.

That said, the adoption of such pairing primitives increases the expressive power of the framework. While this move does not enrich the ontological commitments of the theory—no new entities are posited—it does raise what is sometimes called the *ideological commitment* of the language, i.e., the amount of primitive concepts assumed in its formulation.[59] I discuss the trade-off between ontological parsimony and ideological richness in more detail in Section 6.2.1.

Note that the above definition postulates a sequence of higher-order variables $(X^1, X^2, ..., X^n)$, each of which satisfies the model predicate:

Lemma 3.36. Let X^{n+1} be such that $\mathbb{M}(X^{n+1})$, and $X^1, X^2, ..., X^n$ be as postulated by definition 3.34. Then, for all $s\in\{1,...,n\}$, $\mathbb{M}(X^s)$.

Proof. Let us first observe that $\sqsubseteq$ is transitive: Let X^i, X^j, X^k be such that $X^i\sqsubseteq X^j$ and $X^j\sqsubseteq X^k$, and let x^l be an arbitrary element in X^i, i.e., $X^i(x^l)$. Then, by $X^i\sqsubseteq X^j$, it follows that $X^j(x^l)$, and by $X^j\sqsubseteq X^k$, it follows that $X^k(x^l)$.

Now, assume for *reductio* that $\mathbb{M}(X^{n+1})$ holds, but for some $1\leq s\leq n$ and X^s from the sequence $X^1,...,X^n$ postulated by definition 3.34, $\neg\mathbb{M}(X^s)$. There are two possible cases for $\neg\mathbb{M}(X^s)$:

Case 1: X^s is empty, but this is ruled out by definition 3.34.

Case 2: There is an element y^i for $0\leq i<s$ such that $X^s(y^i)$, but which is neither of the form $y^i=\langle\ulcorner\forall\urcorner,x^i\rangle$, nor of the form $y^i=\langle x^i,z^i\rangle$. But then, by transitivity of $\sqsubseteq$, we have $X^{n+1}(y^i)$, and consequently, $\neg\mathbb{M}(X^{n+1})$, contradicting the assumption. □

Let me conclude this section with a final remark about $\mathbb{M}(X^{n+1})$: as I have defined it, $\mathbb{M}(X^{n+1})$ applies not only to free higher-order variables, but also to models in the classical sense of model theory. More precisely,

59 See Cowling [23].

given a classical, model-theoretic model for a higher-order object language $\mathcal{M}^n = \langle M, \mathcal{P}(M), ..., \mathcal{P}^{n-1}(M), I^n \rangle$, we can always transform this into an RU-model $\mathbb{M}(X^{n+1})$. In fact, this is what makes the RU-approach a strict generalization of the classical, model-theoretic approach:

Theorem 3.37. For every classical $\mathcal{L}^n$-model $\mathcal{M}^n = \langle M, \mathcal{P}(M), ..., \mathcal{P}^{n-1}(M), I^n \rangle$, there is a corresponding RU-model $\mathbb{M}(X^n)$.

Proof. Let $\mathcal{M}^n = \langle M, ..., \mathcal{P}^{n-1}(M), I^n \rangle$ be a classical $\mathcal{L}^n$-structure. There exists a sequence of languages $\mathcal{L}^1 \subseteq \mathcal{L}^2 \subseteq ... \subseteq \mathcal{L}^{n-1}$ and interpretations $I^1 \subseteq I^2 \subseteq ... \subseteq I^{n-1}$. Assume further, that σ^n is an $\mathcal{L}^n$-assignment.

Take as X^1

- all $\langle \ulcorner \forall \urcorner, x^0 \rangle$ for $x^0 \in M$;
- all $\langle x^0, y^0 \rangle$ for x^0 a constant or variable type-0 term and

$$y^0 = \begin{cases} I^n(x^0) & \text{if } x^0 \text{ is a constant,} \\ \sigma^n(x^0) & \text{if } x^0 \text{ is a variable.} \end{cases}$$

Take as X^2

- all $\langle \ulcorner \forall \urcorner, x^1 \rangle$ for $x^1 \in \mathcal{P}(M)$;
- all $\langle x^1, \langle y_1^0, ..., y_l^0 \rangle \rangle$ for x^1 a constant or variable type-1 term, and for any $s \in \{1, ..., l\}$

$$y_s^0 = \begin{cases} I^n(x^0) & \text{for some constant type-0 term } x^0, \\ \sigma^n(x^0) & \text{for some variable type-0 term } x^0, \end{cases}$$

and $\langle y_1^0, ..., y_l^0 \rangle \in I^n(x^1)$.

$\vdots$

Take as X^n

- all $\langle \ulcorner \forall \urcorner, x^{n-1} \rangle$ for $x^{n-1} \in \mathcal{P}^{n-1}(M)$;
- all $\langle x^{n-1}, \langle y_1^i, ..., y_l^j \rangle \rangle$ for x^{n-1} a constant or variable type-n term, and for any $s \in \{1, ..., l\}$, any $i, j < n-1$

$$y_s^{i/j} = \begin{cases} I^n(x^{i/j}) & \text{for some constant type-}i/j \text{ term } x^{i/j}, \\ \sigma^n(x^{i/j}) & \text{for some variable type-}i/j \text{ term } x^{i/j}. \end{cases}$$

and $\langle y_1^i, ..., y_l^j \rangle \in I^n(x^{n-1})$.

The resulting object is non-empty, all its elements satisfy the interpretational objects requirement, and it contains interpretations for all $\mathcal{L}^n$-expression. □

The other direction breaks down in the case of absolutist models for the simple reason that, for an RU-model $\mathbb{M}(X^n)$ with an absolutely unrestricted quantifier domain, there is no set that can be taken as the (first-order) domain. However, when we weaken the assumptions to only set-sized RU-models, theorem 3.37 can actually be turned into a biconditional:

Theorem 3.38. For every $\mathcal{L}^n$-RU-model $\mathbb{M}(X^n)$ with a set-sized domain, there is a corresponding classical $\mathcal{L}^n$-model $\mathcal{M}^n = \langle M, \mathcal{P}(M), ..., \mathcal{P}^{n-1}(M), I^n \rangle$.

Proof. Reverse the construction from the proof of theorem 3.37: for any $s \in \{0, ..., n\}$, take all x^s such that $X^{s+1}(\langle \ulcorner \forall \urcorner, x^s \rangle)$ as the domain of s-order quantification. Moreover, construct the interpretation function I^n such that $\langle y_1^i, ..., y_l^j \rangle \in I^n(x^s)$ for any $y_1^i, ..., y_l^j$ such that $X^{s+1}(\langle x^s, \langle y_1^i, ..., y_l^j \rangle \rangle)$. □

3.4.3 *Domain and Variation*

The next step is to extract the domain component of the RU-model. The goal is to isolate all and only those objects of a certain type that can be quantified over. I will use the domain component to define a variation of a model, which serves as the higher-order analogue of a variation of an assignment. (Recall that definition 3.34 already incorporates the notion of a variable assignment. Therefore, I will define the variation of a whole model, rather than that of a variable assignment.)

Definition 3.39. Let $0 \leq s \leq n+1$. For all variable type-$(s+1)$ and type-$(n+1)$ terms Y^{s+1} and X^{n+1}, the $\mathcal{L}^{n+1}$-formula "Y^{s+1} is the type-s RU-domain of the RU-model X^{n+1}", in symbols $\mathbb{D}(X^{n+1}, Y^{s+1})$, is defined as:

$$\mathbb{D}(X^{n+1}, Y^{s+1}) :\leftrightarrow \mathbb{M}(X^{n+1}) \wedge \forall x^s \Big(Y^{s+1}(x^s) \leftrightarrow \exists y^s (X^{s+1}(y^s) \wedge y^s = \langle \ulcorner \forall \urcorner, x^s \rangle) \Big)$$

As intended, the above definition isolates all and only those entities corresponding to type-s expressions of $\mathbb{M}(X^{n+1})$ that can be quantified over. Thus, when an RU-model $\mathbb{M}(X^n)$ contains interpretational objects of the form $\langle x^s, \langle y_1^i, ..., y_l^j \rangle \rangle$—meaning that in $\mathbb{M}(X^n)$, $x^s(y_1^i, ..., y_l^j)$ holds—there is an object corresponding to the expression x^s which resides in the type-s layer of the domain. Note that the first conjunct of definition 3.39 refers to $\mathbb{M}(X^{n+1})$, and by definition 3.34, X^{s+1} exists. Therefore, X^{s+1} applies to all and only those objects x^s such that X^{s+1} applies to an ordered pair of the form $y^s = \langle \ulcorner \forall \urcorner, x^s \rangle$.

Let me now proceed to the definition of a variant of a model. I will first present the definition and then provide an explanation afterward:

Definition 3.40. Let $0 \leq s \leq n$. For any type-s variable q^s, and any type-$(n+1)$ variables X^{n+1}, Y^{n+1}, the $\mathcal{L}^{n+1}$-formula "the RU-model Y^{n+1} is a RU-q^s-variant of the RU-model X^{n+1}", in symbols $\mathbb{V}(q^s, X^{n+1}, Y^{n+1})$, is defined as:

$$\mathbb{V}(q^s, X^{n+1}, Y^{n+1}) :\leftrightarrow \mathbb{M}(X^{n+1}) \wedge \mathbb{M}(Y^{n+1}) \wedge \exists Z^1 (\mathbb{D}(X^1, Z^1) \wedge \mathbb{D}(Y^1, Z^1)) \wedge ... \wedge$$

$$\exists Z^{n+1}(\mathbb{D}(X^{n+1}, Z^{n+1}) \wedge \mathbb{D}(Y^{n+1}, Z^{n+1})) \wedge$$
$$\forall x^0 \forall z^0 \Big(X^1(\langle x^0, z^0 \rangle) \leftrightarrow Y^1(\langle x^0, z^0 \rangle) \Big) \wedge ... \wedge$$
$$\forall x^s \Big[x^s \neq q^s \rightarrow \forall z^s \Big(X^{s+1}(\langle x^s, z^s \rangle) \leftrightarrow Y^{s+1}(\langle x^s, z^s \rangle) \Big) \Big] \wedge ... \wedge$$
$$\forall x^n \forall z^n \Big(X^{n+1}(\langle x^n, z^n \rangle) \leftrightarrow Y^{n+1}(\langle x^n, z^n \rangle) \Big)$$

According to definition 3.40, Y^{n+1} is an RU-q^s-variant of X^{n+1} if and only if the following conditions hold: First, both X^{n+1} and Y^{n+1} are RU-models. Second, there exists a sequence $Z^1, ..., Z^{n+1}$ such that X^1 and Y^1 share the same domain Z^1, ..., and X^{n+1} and Y^{n+1} share the same domain Z^{n+1} (note that, by definition 3.34, the sequences $X^1, ..., X^n, Y^1, ..., Y^n$ exist, and by lemma 3.36, $\mathbb{M}(X^1), ..., \mathbb{M}(X^n)$ as well as $\mathbb{M}(Y^1), ..., \mathbb{M}(Y^n)$ also hold). Third, the models X^{n+1} and Y^{n+1} agree on the interpretation of all expressions of order i for $i \neq s$, i.e., they agree on all expressions that are *not* of order s. Finally, X^{n+1} and Y^{n+1} agree on the interpretation of all type-s expressions except for q^s. In other words, X^{n+1} and Y^{n+1} may disagree on the interpretation of q^s but only on that one expression. Consequently, if Y^{n+1} is an RU-variant of X^{n+1}, this means that both are RU-models and agree on the interpretation of all expressions—that is, on all (constant or variable) terms of every type from 0 to n—except possibly on the interpretation of q^s. Note that this also makes X^{n+1} an RU-variant of itself, since the condition only specifies that the two models agree on all expressions except for q^s.

3.4.4 *Satisfaction*

Having introduced the notions of an RU-model, the RU-domain component, and RU-variation, we can now define satisfaction for RU-models. Typically, satisfaction is defined as a relation between a model, an assignment, and a (set of) formula(e). However, since I have defined variable assignments to be part of the model itself, satisfaction here is defined as a relation between a model and a formula. This is primarily a matter of notational convention rather than a substantive conceptual difference. The underlying idea remains the same: if $\varphi(x_1, \ldots, x_n)$ is a formula with all its free variables displayed (ignoring the types of variables for now), and if $\mathbb{M}(X^n)$ is an RU-model that includes an assignment of variables, then $\varphi(x_1, \ldots, x_n)$ is true relative to $\mathbb{M}(X^n)$ if and only if the assignment maps $x_1, \ldots, x_n$ to objects $a_1, \ldots, a_n$, respectively, such that φ holds of these objects $a_1, \ldots, a_n$.

There are several distinct notions of satisfaction, two (or three, depending on how one counts) of which will be important here: the classical Tarskian account of satisfaction,[60] as well as the non-classical (strong and weak) Kleene accounts of satisfaction.[61]

Tarskian satisfaction is grounded in classical, two-valued logic and validates the principle of bivalence, according to which every sentence of the underlying language has exactly one of the two truth values: true or false. In particu-

60 See Tarski [156] (1933 in Polish).
61 See Kleene [78].

lar, this means that for any relation (of arbitrary arity) z^s, and for objects $t_1^i, \ldots, t_m^j$ of the appropriate types (i.e., $i, j < s$), exactly one of $z^s(t_1^i, \ldots, t_m^j)$ or $\neg z^s(t_1^i, \ldots, t_m^j)$ holds. Moreover, for any predicate in the language, the extension of the predicate (the collection of objects for which the predicate holds) and the anti-extension of the predicate (the collection of objects for which its negation holds) jointly exhaust the domain.

Kleene satisfaction generalizes Tarskian satisfaction in that it allows for the possibility of truth-value gaps. That is, there may be predicates (or relations) and objects in the domain such that the object is neither in the extension nor in the anti-extension of the predicate. This situation is modeled by introducing a third truth value, $1/2$. If an atomic sentence $z^s(k^i)$ has truth value $1/2$, this means that it is undefined whether k^i possesses the property z^s or not.

This implies that the usual connection between a sentence not being true and the truth of its negation breaks down. In classical logic, due to the principle of bivalence, if φ does not hold, then $\neg\varphi$ holds, i.e., if φ has truth value 0, then $\neg\varphi$ has truth value 1. This connection, however, fails in the case of Kleene satisfaction. If φ is not true, this can mean that φ has truth value 0 or truth value $1/2$. Yet, only if φ has truth value 0, $\neg\varphi$ has truth value 1. If, however, φ has truth value $1/2$, then so does $\neg\varphi$. Thus, if a sentence is not true, it does *not* automatically follow that its negation is true.

Regarding the evaluation of complex formulae, both strong and weak Kleene satisfaction retain a fully classical fragment. If φ and ψ are formulae of the underlying language, both having a classical truth value, then the truth values of $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \to \psi$, $\exists x\, \varphi$, and $\forall x\, \varphi$ are computed exactly as in the classical case. Only when at least one of the formulae involved has the non-classical truth value $1/2$ can the complex formulae also take the value $1/2$. In the case of strong Kleene satisfaction, even if one subformula has the value $1/2$, the truth value of the complex formula can still be calculated following the classical truth tables. More generally, let $\mathfrak{v}$ be a valuation, mapping sentences to the truth values $\{0, 1/2, 1\}$, then the truth values of complex formulae according to strong Kleene evaluation are calculated as follows:

$$\begin{aligned}
\mathfrak{v}(\neg\varphi) &= 1 - \mathfrak{v}(\varphi) \\
\mathfrak{v}(\varphi \wedge \psi) &= \min\{\mathfrak{v}(\varphi), \mathfrak{v}(\psi)\} \\
\mathfrak{v}(\varphi \vee \psi) &= \max\{\mathfrak{v}(\varphi), \mathfrak{v}(\psi)\} \\
\mathfrak{v}(\varphi \to \psi) &= \max\{1 - \mathfrak{v}(\varphi), \mathfrak{v}(\psi)\} \\
\mathfrak{v}(\forall x\varphi) &= \min\{\mathfrak{v}'(\varphi) \mid \mathfrak{v}' \text{ an } x\text{-variant of } \mathfrak{v}\} \\
\mathfrak{v}(\exists x\varphi) &= \max\{\mathfrak{v}'(\varphi) \mid \mathfrak{v}' \text{ an } x\text{-variant of } \mathfrak{v}\}
\end{aligned}$$

Consequently, a disjunction is true if and only if at least one of the disjuncts is true. However, if one disjunct has truth value 0 and the other has truth value $1/2$, then the disjunction itself takes the value $1/2$. Similarly, in the case of conjunction, which can have the truth value $1/2$ if one conjunct is true and the other has value $1/2$, or if both conjuncts have the value $1/2$.

In the case of *weak* Kleene satisfaction, the truth value of complex formulae coincides with their classical evaluation as long as all subformulae have classical truth values. However, if at least one subformula takes the non-classical value

$1/2$, then the entire complex formula also takes the value $1/2$.[62] This behavior can be summarized as follows:

$$\mathfrak{v}(\neg\varphi) = \begin{cases} 1 - \mathfrak{v}(\varphi), & \text{if } \mathfrak{v}(\varphi) \neq 1/2 \\ 1/2, & \text{otherwise} \end{cases}$$

$$\mathfrak{v}(\varphi \wedge \psi) = \begin{cases} \min\{\mathfrak{v}(\varphi), \mathfrak{v}(\psi)\}, & \text{if } \mathfrak{v}(\varphi) \neq 1/2 \text{ and } \mathfrak{v}(\psi) \neq 1/2 \\ 1/2, & \text{otherwise} \end{cases}$$

$$\mathfrak{v}(\varphi \vee \psi) = \begin{cases} \max\{\mathfrak{v}(\varphi), \mathfrak{v}(\psi)\}, & \text{if } \mathfrak{v}(\varphi) \neq 1/2 \text{ and } \mathfrak{v}(\psi) \neq 1/2 \\ 1/2, & \text{otherwise} \end{cases}$$

$$\mathfrak{v}(\varphi \to \psi) = \begin{cases} \max\{1 - \mathfrak{v}(\varphi), \mathfrak{v}(\psi)\}, & \text{if } \mathfrak{v}(\varphi) \neq 1/2 \text{ and } \mathfrak{v}(\psi) \neq 1/2 \\ 1/2, & \text{otherwise} \end{cases}$$

$$\mathfrak{v}(\forall x\varphi) = \begin{cases} \min\{\mathfrak{v}'(\varphi) \mid \mathfrak{v}' \text{ an } x\text{-variant of } \mathfrak{v}\}, & \text{if } \mathfrak{v}'(\varphi) \neq 1/2 \\ & \text{for all } \mathfrak{v}' \\ 1/2, & \text{otherwise} \end{cases}$$

$$\mathfrak{v}(\exists x\varphi) = \begin{cases} \max\{\mathfrak{v}'(\varphi) \mid \mathfrak{v}' \text{ an } x\text{-variant of } \mathfrak{v}\}, & \text{if } \mathfrak{v}'(\varphi) \neq 1/2 \\ & \text{for all } \mathfrak{v}' \\ 1/2, & \text{otherwise} \end{cases}$$

If we restrict the set of logical connectives to $\{\neg, \vee, \forall\}$, the main difference between strong and weak Kleene logic concerns the disjunction case. In strong Kleene logic, $\varphi \vee \psi$ is true if and only if at least one of the disjuncts is true. By contrast, in weak Kleene logic, $\varphi \vee \psi$ is true if and only if one disjunct is true *and* both disjuncts have a classical truth value.

Different intuitions might motivate these distinct treatments of disjunction. For instance, one might think that a disjunction such as $\varphi \vee 0 = 0$ should be true, no matter what. Alternatively, one might prefer the *off-topic* interpretation developed by Beall [6]. In this work, I will employ both approaches: I will use strong Kleene logic in chapter 4 and weak Kleene logic in chapter 5, where I will also explain my reasons for doing so.

Let me end with a brief remark on the relation between Kleene and Tarski satisfaction. Since, according to Kleene satisfaction, the truth value of complex formulae is compositionally determined by the truth values of their parts—just as in classical logic—and since atomic formulae are the minimal parts of complex formulae, a complex formula can take the value $1/2$ only if it contains an atomic subformula with that value. This can occur either because some predicates have gaps or because some terms fail to denote. However, in the limiting case where all terms denote and all atomic predicates behave

62 Typically, in weak Kleene logic, the third truth value is interpreted as *nonsensical* or *meaningless*, and is sometimes referred to as *infectious*, because as soon as one subformula is nonsensical, the whole complex formula becomes *infected with meaninglessness*; see, e.g., Beall and Van Fraassen [8] and Bochvar and Bergmann [11].

classically (i.e., where the extension and antiextension of each predicate exhaust the domain, so that every object belongs to exactly one of them), Kleene and Tarski satisfaction coincide.

Formally, let $\models$ denote the classical satisfaction relation, and let $\models_{sk}$ and $\models_{wk}$ denote strong and weak Kleene satisfaction relations, respectively, all defined over the same language $\mathcal{L}$. If we assume that all terms of $\mathcal{L}$ denote, and that for every atomic predicate and relation, the extension and antiextension jointly exhaust the domain, then for every set of $\mathcal{L}$-formulae Γ and every $\mathcal{L}$-formula φ, the following are equivalent:

1. $\Gamma \models \varphi$
2. $\Gamma \models_{wk} \varphi$
3. $\Gamma \models_{sk} \varphi$

This can be proved through what is known as *classical recapture*.[63] When working within a three-valued logic such as Kleene's systems, we can always recapture classical logic by adopting classical principles in a way that ensures the relevant sentences receive a classical truth value.

For example, in the case of strong Kleene logic, the logical connectives are defined so that they yield an indeterminate value only when necessary—such as when one disjunct is false and the other is indeterminate. In such cases, we can force determinacy by adopting instances of *double negation elimination* (DNE). The rationale behind this is as follows: if φ were indeterminate, then both $\neg\varphi$ and $\neg\neg\varphi$ would also be indeterminate, and thus the instance of DNE would fail. Therefore, assuming DNE for φ forces φ to have a classical truth value. In this sense, DNE allows us to *recapture* classical behavior within strong Kleene logic.

A parallel situation arises in weak Kleene logic when we consider the *law of excluded middle* (LEM). Recall that in weak Kleene logic, compound formulae are highly sensitive to indeterminacy: if even a single subformula is indeterminate, the entire formula becomes indeterminate. By assuming an instance of LEM for a sentence φ, we again force φ to assume a classical truth value.

With these observations in place, I will now proceed to prove the classical recapture result for both strong and weak Kleene satisfaction:

Theorem 3.41. Let $\mathcal{L}^n$ be an n^{th}-order language, let $\models_{sk}$ and $\models_{wk}$ denote the satisfaction relations for the strong and weak Kleene semantics respectively, while $\models$ denotes classical satisfaction. Let φ be a sentence of $\mathcal{L}^n$. Then:

1. If $\models_{sk} (\neg\neg\varphi \leftrightarrow \varphi)$, then for every $\mathcal{L}^n$-structure $\mathcal{M}^n$,

$$\mathcal{M}^n \models_{sk} \varphi \quad \Longleftrightarrow \quad \mathcal{M}^n \models \varphi.$$

2. If $\models_{wk} (\varphi \vee \neg\varphi)$, then for every $\mathcal{L}^n$-structure $\mathcal{M}^n$,

$$\mathcal{M}^n \models_{wk} \varphi \quad \Longleftrightarrow \quad \mathcal{M}^n \models \varphi.$$

63 See, for instance, Rosenblatt [141].

Proof. By induction on the complexity of φ.

1. Strong Kleene Case. Assume that

   $$\models_{sk} (\neg\neg\varphi \leftrightarrow \varphi)$$

   holds. If φ were indeterminate, then so would $\neg\varphi$ and hence $\neg\neg\varphi$, causing the equivalence $\neg\neg\varphi \leftrightarrow \varphi$ to fail. Thus, the assumption forces φ to be *determinate*. For atomic formulae, the strong Kleene evaluation coincides with classical truth assignments. By the induction hypothesis, every proper subformula of φ is evaluated classically, and the strong Kleene evaluation (when restricted to classical truth values) agrees with the classical ones. Hence, for every $\mathcal{L}^n$-structure $\mathcal{M}^n$, we have

   $$\mathcal{M}^n \models_{sk} \varphi \quad \Longleftrightarrow \quad \mathcal{M}^n \models \varphi.$$

2. Weak Kleene Case. Assume that

   $$\models_{wk} (\varphi \vee \neg\varphi)$$

   holds. In weak Kleene semantics the disjunction $\varphi \vee \neg\varphi$ is true only when φ is bivalent—that is, when φ takes either the value true or false, ruling out indeterminacy. This assumption ensures that every subformula of φ is determinate. Since, for atomic formulae, weak Kleene evaluation agrees with classical evaluation, and by the induction hypothesis each compound formula built from determinate subformulae behaves classically, it follows that for every $\mathcal{L}^n$-structure $\mathcal{M}^n$,

   $$\mathcal{M}^n \models_{wk} \varphi \quad \Longleftrightarrow \quad \mathcal{M}^n \models \varphi.$$

□

Let me now define classical Tarskian satisfaction (following Rayo and Uzquiano [135] and Rayo and Williamson [136]), strong Kleene satisfaction (following Rossi [144]), and weak Kleene satisfaction for the higher-order semantics. As mentioned earlier, all these notions will be employed in the subsequent chapters. I begin with Tarskian satisfaction.

The aim is to define, for a given object language of type n, a satisfaction predicate $\mathsf{TSat}^n(\varphi, X^{n+1}, Y^{n+1})$, formulated in a type-$(n+1)$ metalanguage, such that $\mathsf{TSat}^n(\varphi, X^{n+1}, Y^{n+1})$ mirrors the standard model-theoretic definition of satisfaction, $\models$, but remains compatible with the RU approach and, consequently, with generality absolutism. (To abuse notation slightly, this would mean that $\mathsf{TSat}^n(\varphi, X^{n+1}, Y^{n+1})$ holds if and only if $\mathbb{M}(X^{n+1}) \models \varphi$.) This is achieved by the following definition, which will be followed by an explanatory remark:

Definition 3.42 (Tarski Satisfaction)**.** Let $\mathcal{L}^n$ be a type-n language, let i, j, s, n be given such that $0 \leq i \leq j < s \leq n$, let X^{n+1} and Y^{n+1} be type-$(n+1)$ variables, and let φ be an $\mathcal{L}^n$-sentence. The $\mathcal{L}^{n+1}$-formula "the RU-model X^{n+1} with RU-domain Y^{n+1} Tarski satisfies φ", in symbols $\mathsf{TSat}^n(\varphi, X^{n+1}, Y^{n+1})$, is defined as: $\mathsf{TSat}^n(\varphi, X^{n+1}, Y^{n+1})$:iff

1. $\mathbb{M}(X^{n+1})$ and $\mathbb{D}(X^{n+1}, Y^{n+1})$ and $\varphi \in \mathsf{Sent}_{\mathcal{L}^n}$ and
2. $\varphi \equiv z^s(t^i_1, ..., t^j_m)$ and $X^{s+1}(\langle z^s, \langle t^i_1, ..., t^j_m\rangle\rangle)$, or
3. $\varphi \equiv \neg\psi$ and $\neg\mathsf{TSat}^n(\psi, X^{n+1}, Y^{n+1})$, or
4. $\varphi \equiv \psi \vee \chi$ and $\mathsf{TSat}^n(\psi, X^{n+1}, Y^{n+1})$ or $\mathsf{TSat}^n(\chi, X^{n+1}, Y^{n+1})$, or
5. $\varphi \equiv \forall z^s \psi(z^s)$ for $\psi \in \mathsf{For}_{\mathcal{L}^s}$, and for every W^{n+1} s.t. $\mathbb{V}(z^s, W^{n+1}, X^{n+1})$, $\mathsf{TSat}^n(\psi(z^s), W^{n+1}, Y^{n+1})$.

TSat inductively defines the extension of the satisfaction predicate for an n^{th}-order object language in an $(n+1)^{\text{th}}$-order metalanguage. According to this definition, TSat^n holds of φ, X^{n+1}, and Y^{n+1} if and only if:

1. φ is an $\mathcal{L}^n$-sentence;
2. X^{n+1} is an RU-model with RU-domain Y^{n+1}; and
3. φ holds in X^{n+1} according to the standard recursive satisfaction clauses (clauses 2–5).

Next, I introduce the definitions of weak and strong Kleene satisfaction within this higher-order framework. As in the Tarskian case, the goal is to define a predicate for weak Kleene satisfaction first (the definition of strong Kleene satisfaction will require only minor modifications). Thus, $\mathsf{WKSat}^n(\varphi, X^{n+1}, Y^{n+1})$ is a predicate of a type-$(n+1)$ metalanguage, such that $\mathsf{WKSat}^n(\varphi, X^{n+1}, Y^{n+1})$ parallels the standard model-theoretic definition of $\models_{wk}$, while remaining compatible with the RU framework and generality absolutism. Below, I present both definitions, followed by an explanatory remark:

Definition 3.43 (Weak Kleene Satisfaction)**.** Let $\mathcal{L}^n$ be a type-n language, let i, j, s, n be given such that $0 \leq i \leq j < s \leq n$, let X^{n+1} and Y^{n+1} be type-$(n+1)$ variables, and let φ be an $\mathcal{L}^n$-sentence. The $\mathcal{L}^{n+1}$-formula "the RU-model X^{n+1} with RU-domain Y^{n+1} weakly Kleene satisfies φ", in symbols $\mathsf{WKSat}^n(\varphi, X^{n+1}, Y^{n+1})$, is defined as: $\mathsf{WKSat}^n(\varphi, X^{n+1}, Y^{n+1})$:iff

1. $\mathbb{M}(X^{n+1})$ and $\mathbb{D}(X^{n+1}, Y^{n+1})$ and $\varphi \in \mathsf{Sent}_{\mathcal{L}^n}$ and
2. $\varphi \equiv z^s(t^i_1, ..., t^j_m)$ and $X^{s+1}(\langle z^s, \langle t^i_1, ..., t^j_m\rangle\rangle)$, or
3. $\varphi \equiv \neg z^s(t^i_1, ..., t^j_m)$ and $\neg X^{s+1}(\langle z^s, \langle t^i_1, ..., t^j_m\rangle\rangle)$, or
4. $\varphi \equiv \neg\neg\psi$ and $\mathsf{WKSat}^n(\psi, X^{n+1}, Y^{n+1})$, or
5. $\varphi \equiv \psi \vee \chi$ and $\mathsf{WKSat}^n(\psi, X^{n+1}, Y^{n+1})$ or $\mathsf{WKSat}^n(\chi, X^{n+1}, Y^{n+1})$, or $\mathsf{WKSat}^n(\neg\psi, X^{n+1}, Y^{n+1})$ or $\mathsf{WKSat}^n(\chi, X^{n+1}, Y^{n+1})$, or $\mathsf{WKSat}^n(\psi, X^{n+1}, Y^{n+1})$ or $\mathsf{WKSat}^n(\neg\chi, X^{n+1}, Y^{n+1})$, or
6. $\varphi \equiv \neg(\psi \vee \chi)$ and $\mathsf{WKSat}^n(\neg\psi, X^{n+1}, Y^{n+1})$ and $\mathsf{WKSat}^n(\neg\chi, X^{n+1}, Y^{n+1})$, or
7. $\varphi \equiv \forall z^s \psi(z^s)$ for $\psi \in \mathsf{For}_{\mathcal{L}^s}$, and for every W^{n+1} s.t. $\mathbb{V}(z^s, W^{n+1}, X^{n+1})$, $\mathsf{WKSat}^n(\psi(z^s), W^{n+1}, Y^{n+1})$, or

8. $\varphi \equiv \neg\forall z^s \psi(z^s)$ for $\psi \in \mathsf{For}_{\mathcal{L}^s}$, and for some W s.t. $\mathbb{V}(z^s, W^{n+1}, X^{n+1})$, $\mathsf{WKSat}^n(\neg\psi(z^s), W^{n+1}, Y^{n+1})$.

Definition 3.44 (Strong Kleene Satisfaction)**.** Let $\mathcal{L}^n$ be a type-n language, let i, j, s, n be given such that $0 \leq i \leq j < s \leq n$, let X^{n+1} and Y^{n+1} be type-$(n+1)$ variables, and let φ be an $\mathcal{L}^n$-sentence. The $\mathcal{L}^{n+1}$-formula "the RU-model X^{n+1} with RU-domain Y^{n+1} strongly Kleene satisfies φ", in symbols $\mathsf{SKSat}^n(\varphi, X^{n+1}, Y^{n+1})$, is defined as: $\mathsf{SKSat}^n(\varphi, X^{n+1}, Y^{n+1})$:iff

1. $\mathbb{M}(X^{n+1})$ and $\mathbb{D}(X^{n+1}, Y^{n+1})$ and $\varphi \in \mathsf{Sent}_{\mathcal{L}^n}$ and
2. $\varphi \equiv z^s(t_1^i, ..., t_m^j)$ and $X^{s+1}(\langle z^s, \langle t_1^i, ..., t_m^j\rangle\rangle)$, or
3. $\varphi \equiv \neg z^s(t_1^i, ..., t_m^j)$ and $\neg X^{s+1}(\langle z^s, \langle t_1^i, ..., t_m^j\rangle\rangle)$, or
4. $\varphi \equiv \neg\neg\psi$ and $\mathsf{SKSat}^n(\psi, X^{n+1}, Y^{n+1})$, or
5. $\varphi \equiv \psi \vee \chi$ and $\mathsf{SKSat}^n(\psi, X^{n+1}, Y^{n+1})$ or $\mathsf{SKSat}^n(\chi, X^{n+1}, Y^{n+1})$, or
6. $\varphi \equiv \neg(\psi \vee \chi)$ and $\mathsf{SKSat}^n(\neg\psi, X^{n+1}, Y^{n+1})$ and $\mathsf{SKSat}^n(\neg\chi, X^{n+1}, Y^{n+1})$, or
7. $\varphi \equiv \forall z^s \psi(z^s)$ for $\psi \in \mathsf{For}_{\mathcal{L}^s}$, and for every W^{n+1} s.t. $\mathbb{V}(z^s, W^{n+1}, X^{n+1})$, $\mathsf{SKSat}^n(\psi(z^s), W^{n+1}, Y^{n+1})$, or
8. $\varphi \equiv \neg\forall z^s \psi(z^s)$ for $\psi \in \mathsf{For}_{\mathcal{L}^s}$, and for some W s.t. $\mathbb{V}(z^s, W^{n+1}, X^{n+1})$, $\mathsf{SKSat}^n(\neg\psi(z^s), W^{n+1}, Y^{n+1})$.

As in the case of TSat, the predicates WKSat and SKSat define the extension of (weak and strong) Kleene satisfaction for a type-n object language in a type-$(n+1)$ metalanguage. The basic requirements are analogous: X^{n+1} must be an RU-model with RU-domain Y^{n+1}, and φ an $\mathcal{L}^n$-sentence that holds in X^{n+1} according to appropriate recursive clauses. However, unlike TSat, WKSat and SKSat cannot appeal to the classical negation clause (that φ is true iff $\neg\varphi$ is not true), for reasons discussed earlier.

The main difference between WKSat and SKSat appears in the disjunction clause (5): for WKSat, $\psi \vee \chi$ is true if at least one disjunct is true and the other has a classical truth value (the definition above lists all cases). In contrast, for SKSat it suffices that one disjunct is true, regardless of the other's value.

Although Kleene satisfaction is a three-valued semantics, the definitions above do not presuppose this. What is defined is merely the extension of the satisfaction predicate—that is, the set of sentences true in a model. By contrast, an *antiextension* would collect all sentences false in a model, which we can denote ¬WKSat and ¬SKSat, respectively. Thus, WKSat and SKSat correspond to truth value 1, while ¬WKSat and ¬SKSat correspond to 0. The gap (truth value 1/2) contains precisely those sentences belonging to neither. For instance, if a term a denotes an object that is neither in the extension nor antiextension of some predicate P, then $P(a)$ will be in the gap—as will, e.g., $P(a) \wedge \varphi$ (for $\mathfrak{v}(\varphi) > 0$), or $\varphi \vee \psi$ (for arbitrary ψ under weak Kleene satisfaction), and so on.

Defining only the extension has some advantages. First, it allows us to recover classical behavior when needed.[64] Second, we can easily reconstruct the antiextension (and thus the gap) via:

$$\neg\mathsf{WKSat}^n(\varphi, X^{n+1}, Y^{n+1}) := \{\neg\varphi \mid \mathsf{WKSat}^n(\varphi, X^{n+1}, Y^{n+1})\}$$

and

$$\neg\mathsf{SKSat}^n(\varphi, X^{n+1}, Y^{n+1}) := \{\neg\varphi \mid \mathsf{SKSat}^n(\varphi, X^{n+1}, Y^{n+1})\}.$$

This concludes the presentation of the technical background. In the following chapters, I will apply the RU-approach in the truth- and set-theoretic context. Specifically, I develop absolutist-friendly, higher-order versions of Glanzberg's contextualist truth theory (chapter 4), as well as a categorically axiomatized set theory in the sense of the iterative conception (chapter 5).

64 This will be crucial in later chapters, especially in chapter 4, where I develop a fully classical truth predicate in the sense of Kripke [79].

4

TRUTH-THEORETIC BICONTEXTUALISM FOR HIGHER-ORDER LANGUAGES

In section 2.4, I introduced bicontextualism as a way to combine the best of two worlds: on one hand, we have elegant and sophisticated yet inherently relativistic solutions to the paradoxes; on the other hand, there is the appeal of absolutism. Bicontextualism combines the strengths of both perspectives.

In a truth-theoretic setting, this entails formulating a semantic theory where quantifiers are allowed to range over an unrestricted 'domain', provided the sentence in question is not an expansion sentences (in the sense of section 2.4.1).[1] For expansion sentences, however, the interpretation is necessarily restricted. This aligns with the approach taken in Rossi [144], who develops this position for a first-order object language. In the present chapter, I extend Rossi's approach to accommodate higher-order object languages.

To achieve this generalization, I reconstruct Rossi's adaptation of Glanzberg's contextualist truth theory within the RU framework, but adapted to higher-order languages. This approach ensures that non-expansion sentences are evaluated over an unrestricted 'domain', while expansion sentences are interpreted over a sequence of increasingly large but set-sized domains. This semantic theory upholds relativism specifically for problematic self-referential or liar-like sentences while allowing absolutism for all remaining, unproblematic sentences.

More concretely, I expand the language with a sequence of context-relative truth predicates Tr_α, one for each context c_α, and I construct Kripke fixed-points in two ways: one over the all-inclusive 'domain', and one over the contextually restricted and set-sized domain. This is done for each truth predicate Tr_α. Since there is a sequence of context-relative truth predicates, this involves constructing a sequence of fixed-points, absolutist-friendly as well as relativistic (i.e., contextualist fixed-points, constructed over set-sized yet increasing domains). All sentences of the language (especially the sentences of the truth-free fragment) get evaluated multiple times, and the bicontextualist semantics picks out, for each context c_α, *the right interpretation*. The semantics works such that all the unproblematic sentences (such as $\forall x\ x = x$) get interpreted over the all-inclusive 'domain', while all the expansion sentences like the contextualized liar get interpreted only over restricted domains.

The next step involves defining a bicontextualist relation of logical consequence that selectively applies the appropriate interpretation to each sentence. Formally, I define a relation, $\Gamma \models^n_{bc} \varphi$ (with *bc* for bicontexualism and *n* indicating higher-order languages), in which all non-expansion sentences within $\Gamma \cup \{\varphi\}$ are interpreted over the all-inclusive 'domain', while all expansion sentences are evaluated over set-sized domains. Thus, an argument from Γ to φ

1 Again, 'domain' (with quotation marks) always stands for non-objectual domain, while domain (with quotation marks) stands for an objectual domain in the sense of classical model theory.

will be *bicontextually valid* if and only if, if all sentences in Γ are true either in the absolutist model or in some relativist model, then so is φ.

Before developing this technical framework, I will first provide a motivation for generalizing bicontextualism in this way, i.e., to develop a type-free truth theory for unrestricted higher-order languages. Since absolutism and type-free truth have been discussed in chapter 2 and section 3.2, the current motivation will focus particularly on the higher-order aspects. Specifically, I will argue that a truth theory aiming to capture natural language truth must involve higher-order object languages (section 4.1). Following this motivation, I will develop truth theoretical Bicontextualism for higher-order languages, first heuristically (section 4.2), then I give the technical details (section 4.3). The philosophical implications of higher-order bicontextualism will be discussed in together with the philosophical aspects of set-theoretic bicontextualism (chapter 5) in chapter 6.

4.1 MOTIVATION

The contextualist truth theory developed here is consistent, unlike naïve truth theories which lead to paradox. This point is crucial for understanding the analogy with the set-theoretic part to be discussed in detail in chapter 5: just as the cumulative-iterative conception of sets avoids paradox through stratification, so too does contextualist truth avoid inconsistency via context-parametrization. Thus, the truth-theoretic and set-theoretic applications of bicontextualism are structurally and philosophically aligned.

A key objective of logic is to model natural language, including its structural and semantic complexity. However, certain characteristics of natural language suggest that first-order logic is insufficient for this task. Instead, higher-order logic better models natural language, especially in consideration of principles like categoricity, the non-compactness of some natural language inferences, and Rayo and Linnebo's principle of semantic optimism, which advocates an expansive hierarchy of higher-order languages. Below, I discuss all these points and build an argument that natural language modeling demands not only second-order logic but motivates extending this hierarchy to arbitrary finite orders.

4.1.1 *Natural Language and Categoricity Phenomena*

Mathematical expressions are part of natural language.[2] It is therefore plausible that mathematical phenomena that require full higher-order quantification and appear in natural language support the inclusion of higher-order quantification in our formal semantics for natural language.

Categoricity refers to a theory's ability to uniquely characterize structures up to isomorphism. While first-order logic lacks this ability due to the compactness and Löwenheim–Skolem theorems, full second-order logic can achieve it.[3] For

2 This is not to say that *all* mathematical phenomena occur in natural language. For example, natural language does not permit infinite conjunctions or disjunctions.

3 See section 3.1 for details.

instance, first-order Dedekind-Peano Arithmetic (PA) has many non-isomorphic models, and thus cannot uniquely capture the intended structure of the natural numbers. By contrast, Dedekind showed that second-order arithmetic (PA_2)—which replaces the induction schema in PA with a second-order induction axiom—is categorical: all its models are isomorphic to $\mathbb{N}$.[4] That is, second-order logic can uniquely characterize $\mathbb{N}$ up to isomorphism.

The power of second-order logic to uniquely characterize important mathematical structures extends beyond arithmetic. It allows us to uniquely characterize structures such as $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$.[5] Crucially, this only holds in full second-order logic—where second-order quantifiers range over the entire powerset of the first-order domain.

This has linguistic consequences. When someone says "*The* natural numbers...", the definite article suggests that the speaker intends to refer *at least* to a unique, well-defined structure. (The 'at least' is important for the underlying metaphysical assumption about mathematical objects; I'll come back to it later in this section.) Such utterances—call them *natural language categoricity phenomena*—are best understood as relying on a fundamental capacity of natural language to identify unique structures.

This interpretation, while disputable,[6] is linguistically and cognitively plausible. Speakers seem to succeed, or at least attempt to succeed, in referring uniquely to structures like $\mathbb{N}$, $\mathbb{R}$, or even "the first inaccessible rank of the cumulative hierarchy." If so, natural language quantifiers must function more like those of fully interpreted second-order logic. First-order logic, with its failure to ensure uniqueness, cannot account for this.

Thus, higher-order logic is the best explanation for natural language categoricity phenomena: only by assuming full higher-order quantification can we make sense of natural language's apparent ability to refer uniquely to certain mathematical structures.[7]

This position aligns with what Button and Walsh call the *Infer-to-Stronger Logic Attitude* which infers the need for stronger logical resources from the limitations of first-order logic. Analogously, we may adopt an *Infer-to-Stronger Language Attitude*: if natural language exhibits categoricity phenomena, then our language—not just our logic—must transcend first-order expressive limitations.[8]

Importantly, this is not to claim that natural language quantification *is* higher-order quantification in the formal sense. Natural language and formal logic operate at very different levels of abstraction and theoretical construction. Conflating them—by treating natural language as identical with full second-order semantics—would amount to a category mistake. Natural language is not

4 See Shapiro [149, Sec. 4.2] for the definition of PA_2; see also Button and Walsh [18, Sec. 7.4] and Maddy and Väänänen [98]. All sources provide proofs of Dedekinds categoricity theorem.

5 For set theory, only quasi-categoricity is achievable. See section 3.3.

6 For instance, Hamkins challenges that there is a unique model of arithmetic based on the prediction of a future development of a forcing analogue for arithmetic' [61, p. 428].

7 Categoricity for $\mathbb{N}$ can also be achieved via infinitary logic (e.g., $\mathcal{L}_{\omega_1,\omega}$), and for set theory via stronger logics like $\mathcal{L}_{\kappa^+,\kappa^+}$. But, as mentioned above, infinitary conjunctions and disjunctions are not part of natural language.

8 See Button and Walsh [18, p. 157].

a formal system, and its quantificational mechanisms are not directly reducible to any formal apparatus.

Rather, my claim is more modest. I suggest that there are natural language *categoricity phenomena*: cases in which speakers appear to refer to structures uniquely characterizable up to isomorphism. Expressions like "the natural numbers" are central examples. I assume that such utterances can be successful, i.e., in some cases, speakers successfully refer to the intended structure (e.g., the standard model of arithmetic). This is a substantive but defeasible assumption: I take the use of such expressions at face value, and ask what kinds of semantic resources would be needed to account for that success.

From this explanatory starting point, my argument is conditional: *if* such natural language utterances succeed in uniquely referring to a structure, and *if* we aim to model this success within a semantic theory, then the semantic theory must incorporate resources at least as strong as those found in full higher-order logic. First-order logic lacks the expressive strength to account for this kind of referential uniqueness. In that sense, higher-order quantification (understood in terms of its role within formal semantics) offers the best available explanation for natural language categoricity phenomena. It fills the explanatory gap between the surface features of our language and the model-theoretic behavior of referential expressions.

This explanatory approach does not entail that natural language is formally governed by higher-order logic, nor that all such utterances are to be understood in the sense of a natural language categoricity phenomenon. Rather, it highlights that there are cases in which natural language appears to act *as if* it were supported by strong logical resources—resources that are, in formal settings, captured by higher-order quantification under full semantics.

My account of natural language categoricity phenomena is not committed to a specific metaphysical view about the nature of numbers. As mentioned above, I assume that speakers refer *at least* to a unique structure. In other words, I adopt a neutral *compatibilist* stance: the thesis is compatible with a range of metaphysical perspectives, provided they accept the categoricity of second-order arithmetic. In this way, my focus is primarily epistemological and semantic rather than ontological. I am concerned with how speakers use natural language to successfully refer to mathematical structures, not with what numbers ultimately are.

In particular, structuralist and neo-logicist views both accept that the structure of the natural numbers is unique (up to isomorphism), even though they differ in how they interpret what numbers are. Structuralists emphasize that the identity of mathematical objects is determined by their place in a structure, whereas neo-logicists regard numbers as logical objects defined through abstraction principles, such as Hume's Principle. But both positions accept the claim that there is one natural number *structure*, and that second-order logic—with full semantics—can capture it.

Turning specifically to neo-logicism: part of this position is Frege's Theorem which shows that second-order logic with sufficient comprehension, together

with Hume's Principle entails the axioms of second-order Peano Arithmetic.[9] My position here does not interfere with or contradict this result. In fact, it is orthogonal to it. The neo-logicist derivation yields PA_2, and since PA_2 is categorical, the structure of the natural numbers are unique (up to isomorphism). Therefore, even if the numbers in the neo-logicist sense are logical objects defined through abstraction, the structure they instantiate will align with the one captured by the categoricity theorem.

Neo-logicists usually have a cardinal conception of number (i.e., seeing numbers as measures of equinumerosity classes), whereas structuralists often adopt an ordinal conception (focusing on order type and successor structure). But for the purposes of explaining categoricity phenomena in natural language, this distinction is not fatal. Both approaches, when properly formalized, agree on the relevant structure: the model of second-order arithmetic. Whether one interprets the objects within that model as cardinals or ordinals is not decisive, what matters for my argument is the structure they share.

More specifically, my argument relies only on structural identification of the natural numbers—i.e., classification up to isomorphism—rather than on the stronger notion of identity of mathematical objects. The categoricity of second-order Peano Arithmetic guarantees that any two models of the theory are isomorphic, but not that their objects are *identical*. This reflects an implicit structuralist feature of formalized theories: axioms typically determine a structure only up to isomorphism, and any model satisfying those axioms (such as the set ω of finite von Neumann ordinals or the standard model $\mathbb{N}$) is considered to be equally legitimate in mathematical practice. Crucially, even if one endorses the existence of unique natural numbers—as some realist or abstractionist views do—this is not incompatible with my approach. For the purpose of explaining natural language reference to "the natural numbers," what matters is that *at least* the relevant structure can be picked out, not that it is pinned down up to identity. My focus is thus on the unique structure that speakers intend to refer to, rather than the unique objects that might instantiate it. This is why my view remains fully compatible with theories like neo-logicism: so long as those objects, when formally treated, form a model of second-order arithmetic, they fall within the same isomorphism class that my account targets.[10]

4.1.2 *Non-Compactness in Natural Language*

Griffiths and Paseau [54, 120] have argued that natural language lacks the compactness property characteristic of first-order logic. Compactness implies that if every finite subset of a set of sentences is satisfiable, then the entire set is satisfiable. However, natural language appears to support statements and arguments that lack this property. Griffiths and Paseau's argument is inspired by Boolos [13], and runs as follows: First, observe that the the compactness theorem (i.e., that a set of sentences Γ is satisfiable if and only if each finite

9 The first to notice this seems to be Parsons [115]. For a proof of the Frege's Theorem, see Wright [172]. For more general background on (neo-) logicism, see Hale and Wright [60, 59] and Tennant [157].

10 Thanks to Hosea von Hauff for fruitful discussion on this issue.

subset $\Gamma_{fin} \subseteq \Gamma$ is satisfiable) is equivalent to the following: For a set of sentences Γ and a sentence φ, $\Gamma \models \varphi$ if and only if there is a finite subset $\Gamma_{fin} \subseteq \Gamma$ such that $\Gamma_{fin} \models \varphi$. Now let Γ be the set containing a sentence of the form 'There are at least n planets' for every $n \in \mathbb{N}$, and let φ be the sentence 'There are infinitely many planets'. Clearly, $\Gamma \models \varphi$. However, there is no finite subset $\Gamma_{fin} \subseteq \Gamma$ such that $\Gamma_{fin} \models \varphi$. Since the argument is valid in natural language, its logical consequence relation cannot be compact, demanding instead a logic where compactness does not hold, such as second-order or higher-order logic.

A possible reaction to the argument is that it contains infinitely many premises, and can therefore not be formulated in natural language. Rather, the objection continues, we understand $\Gamma \models \varphi$ only through a finite description of the infinitely many premises in Γ:

> The argument [$\Gamma \models \varphi$] consists of the premises 'There are n planets' for each finite number n and the conclusion 'There are infinitely many planets'. [54, p. 99]

If we understand Γ through a finite proxy, the consequence relation at work satisfies the compactness theorem vacuously, and so it cannot be used to argue against natural language compactness. Griffiths and Paseau respond by pointing out that the important question is whether $\Gamma \models \varphi$ is valid in natural language, and not whether finite beings understand infinite sets only by finite proxies.

Again, a response is possible: If we think about natural language as the totality of utterance of competent speakers, we might have good chances to argue that $\Gamma \models \varphi$ really is *not* part of natural language. But even if this is true, there's another important aspect. If Γ is not part of natural language, we can rephrase $\Gamma \models \varphi$ as follows:

1. The infinitely many premises 'There are n planets'—*one for each finite number n*—together imply the conclusion 'There are infinitely many planets.'

Crucially, 1. invokes phrases like 'infinitely many premises' and 'one for each finite number n.' However, it is a well-known consequence of the compactness theorem that the concept of *arbitrary finitude* cannot be captured in first-order logic.[11] So, even if the objection is correct and can be used to argue against the compactness of natural language consequence, it employs resources which cannot be suitably expressed in a compact consequence relation.

Griffiths and Paseau's argument directly challenges the sufficiency of first-order logic, since non-compactness is not just an incidental feature but rather a core trait of natural language—in particular if we assume that natural language allows to distinguish finiteness from infinity. By ruling out compactness, natural

11 First-order logic can express 'there are exactly/at least n objects' for each n, but not 'there are arbitrarily, but finitely many objects'. [Proof sketch: let Δ be a set of sentences axiomatizing finiteness, and assume that Δ has arbitrary finite models. Let φ_n abbreviate 'there are at least n object', and let $\Gamma = \bigcup_{n \in \mathbb{N}} \varphi_n$. Consider the set $\Delta \cup \Gamma$, and let $\Sigma_n \subseteq \Delta \cup \Gamma$ be a finite subset where n indicates the highest index of $\varphi_n \in \Sigma_n$. Σ_n has a model by assumption. Since n was arbitrary, $\Delta \cup \Gamma$ is finitely satisfiable, and by compactness, it is satisfiable. Let $\mathcal{M}$ be this model. Since $\mathcal{M} \models \Delta$, it models the sentences intended to axiomatize finiteness. However, since $\mathcal{M} \models \Gamma$, it has to be infinite.] This, however, would be crucial for 1. For a philosophical discussion about how the concept of *finitude* is presupposed for first-order model theory, see [18, Ch. 9].

language contains higher-order quantification, as opposed to limited first-order quantification.

4.1.3 *The Principle of Semantic Optimism*

Linnebo and Rayo's [93] *principle of semantic optimism* asserts that for any language $\mathcal{L}$ currently in use, it should always be possible to construct a *generalized semantic theory* (GST) that formalizes all possible interpretations of $\mathcal{L}$—including those that are compatible with absolutism. Semantic optimism thus proposes a framework in which any linguistic or semantic resources we actively employ must be semantically interpretable through a suitable theory.

Moreover, Linnebo and Rayo prove that for an n^{th}-order language $\mathcal{L}^n$, such a GST can only be formulated in a higher-order metalanguage $\mathcal{L}^{n+1}$. This principle suggests that any higher-order language should have a corresponding semantic theory, motivating an ever-expanding hierarchy of logical orders: if we use a second-order language to define the GST of a first-order language, then we already use the second-order language for which we must rely on a third-order language to provide a GST. For similar reasons, we must rely on a fourth-order language to define the GST for a third-order language, and so on.

This hierarchical need is a consequence of the inherent dependence of each language's semantics on a language of a strictly higher order, provided we aim for a semantic theory as strong as GST. Semantic optimism thus demands a framework that inherently supports such a hierarchy to account for all possible finite orders. We are therefore led to posit a series of languages, each increasing in expressive power, to meet the requirements of the principle of semantic optimism.

The implications of semantic optimism are extensive, especially for the foundations of semantics. By ensuring that each language's semantics can be fully interpreted in a higher-order language, semantic optimism posits an open-ended approach to semantics. It resists placing fixed constraints on interpretive frameworks. Instead, by the principle of semantic optimism, we are committed to a system of higher-order languages where each level calls for a new, richer metalanguage to formulate its GST. This framework acknowledges that interpretations are not static or constrained by any given language's resources but require an open-ended model in which each language order provides the resources needed to interpret the previous one.

4.1.4 *Internalizing Semantics through Type-Free Truth*

An important advantage of adopting a type-free truth theory is its ability to handle significantly more important and complex natural language phenomena than the typed counterpart. In a typed theory, truth predicates have limited scope, allowing truth ascriptions only to statements within lower-level languages. Due to this limitation, typed truth makes it impossible to capture natural language phenomena such as blind ascriptions (e.g., 'Everything Kripke says about truth is true') or iterations of the (single and unique) truth predicate (e.g., "P is true' is true'). For instance, modeling blind ascriptions as in the

example would require knowledge about the highest type Kripke ever used to talk about truth. Since this is rather a difficulty than an impossibility for the typed approach, consider the strengthened case of a reciprocal blind ascription. Assume the person who has uttered the Kripke example is A, and assume further that one of Kripke's utterances about truth is 'Everything A says about truth is true'. In that situation, the typed approach comes to its limits: A's utterance has to be of higher type than all of Kripke's utterances, while, simultaneously, Kripke's utterance has to be of higher type than all of A's utterances, which is impossible.[12] By contrast, a type-free truth theory allows the truth predicate to be applied uniformly, without requiring it to ascend in type. In capturing a broader spectrum of natural language truth claims, the type-free approach better models the notion of truth of ordinary discourse.

Within a higher-order framework, a further advantage of this approach is that semantics itself can be internalized within the same framework. This is in line with the general requirements of natural language semantics. Since natural language lacks a truly *external* metalanguage—any metalanguage would still be part of the same overarching language structure—the semantics for natural language have to be formulated within natural language itself. In the higher-order RU framework, interpretations of expressions become higher-order relations that connect linguistic expressions with the world. For example, interpretations of first-order expressions are second-order relations, interpretations of second-order expressions are third-order relations, and so on. This structure, where satisfaction and interpretation are part of the language, aligns with the nature of natural language as being capable of internally defining its own semantics.

4.1.5 *Bringing the Threads Together*

The observations from categoricity, non-compactness, and semantic optimism reinforce each other and form a case for the need of a higher-order framework to model natural language. The categoricity phenomena suggest a need for a minimum of second-order quantifiers to achieve the unique characterizations of important mathematical structures. Non-compactness excludes first-order logic and aligns with the properties of second- and higher-order logic.[13] Finally, the principle of semantic optimism indicates that modeling natural language necessitates a hierarchy of languages, each of which can provide the expressive power for defining the GST of the language below it. Taken together, these points show that natural language is better modeled not merely by second-order logic, but by a system of higher-order logic. Such a system accounts for natural language's capacity to express unique structures, handle non-compact inferences, and contain a semantic depth that calls for an ongoing hierarchy of higher orders.[14]

12 See [34, Ch. 3] and [102, Chs. 3, 4] for a detailed discussion of Kripke's theory of truth.

13 This is not to say that natural language does not contain first-order quantification. The point is that first-order quantification does not exhaust natural language quantification.

14 Further motivations for including higher-order logic might spring from more theoretical observations. For example, by a theorem of Montague, if κ is the least cardinal not characterizable in n^{th}-order logic for $n > 2$, then κ is characterizable in $(n+1)^{\text{th}}$-order logic (see [104, Th. 4]). Furthermore, as Väänänen mentions, there is "a kind of Downward Löwenheim-Skolem

From these considerations, a coherent picture emerges of natural language as formalizable by an open-ended hierarchy of higher-order languages, where each layer provides semantics for the levels beneath. This approach supports a truth theory, where semantics are internalized into the language itself, allowing to define meaning and truth of its own expressions without reliance on external metalanguages. Moreover, this aligns with the type-free approach, where truth and satisfaction are suited to capture the iterative and reciprocal features we encounter in everyday discourse. Such a hierarchy thus provides a powerful and natural model for the semantics of natural language, integrating each level's expressions and truth in a unified, type-free system, while simultaneously making it compatible with generality absolutism.

4.2 HEURISTICS

As argued in section 2.4, the central claim of bicontextualism is that the domain over which a sentence is to be interpreted—whether the unrestricted, all-encompassing 'domain' or a restricted, set-sized domain—depends on the nature of the sentence. Specifically, *non-expansion sentences* are interpreted over the unrestricted 'domain', while *expansion sentences* are restricted to set-sized domains. In this section, I will make a first attempt to develop these ideas in a truth-theoretical setting. This section provides only some heuristic ideas. The full presentation is postponed to the next section.

The distinction between the two types of sentences is rooted in their logical behavior. Expansion sentences, as discussed in chapter 2, are those that cause a domain expansion due to their paradoxical character. This process may involve stepping outside the current set-theoretic universe, guided by reasoning akin to Russell's paradox, or triggering a context shift in a truth-theoretical framework. Such shifts lead to a new context that is more expressive, either by introducing new syntactic expressions, such as a fresh truth predicate, or by adopting a richer ontology of propositions. In contrast, non-expansion sentences do not exhibit such behavior and can thus be evaluated over the all-encompassing 'domain'.

The core idea of bicontextualism is therefore straightforward: interpret all expansion sentences over restricted, set-sized domains and all non-expansion sentences over the unrestricted 'domain'. However, two questions arise: (1) we have only ended with a vague description of expansions sentences at the end of chapter 2. It remains therefore open to give an exact characterization. That is, which sentences are to be classified as expansion sentences, and which are unproblematic in this sense and thus, allow for an unrestricted reading of their quantifiers? (2) following Rossi [144], given that I have identified Parsons-Glanzberg contextualism as a sophisticated, but inherently relativistic solution to the truth-theoretical paradoxes in section 2.2.2, how, then, can we combine contextualism with absolutism for non-expansion sentences?

Theorem for second-order logic", operating at the level of a measurable cardinal [161, sec. 7.2]. So if we want to categorically characterize very big structures, we have to use higher-order logic. And if we want to speak about them as in the case of $\mathbb{N}$, the quantificational resources of natural language crucially have to be of higher-order.

To address this two questions, let us revisit the minimal Kripkean fixed-point. Consider an absolutist-friendly RU model $\mathbb{M}(X)$, which is acceptable in the sense of section 3.2.1 (a translation of acceptability requirements into the RU framework will be given in section 4.3). Ignoring for now the specific order or type of the higher-order variable X, the minimal Kripkean fixed-point constructed over $\mathbb{M}(X)$ contains only grounded sentences. Problematic self-referential sentences, such as the liar or the truth-teller, are excluded. More precisely, constructed over an arithmetical base structure, the minimal fixed-point includes all arithmetical truths, such as sentences of the form $n = n$, $\forall x(x = x)$, and specific cases like $2 + 3 = 5$. Furthermore, for any arithmetical truth φ validated by the standard model $\mathbb{N}$, the fixed-point includes the sentence $\mathsf{Tr}(\lfloor\varphi\rfloor)$ and iterated truth sentences like $\mathsf{Tr}(\lfloor ...\mathsf{Tr}(\lfloor\varphi\rfloor)...\rfloor)$, all of which bottom out at a grounded, truth-free sentence φ. Since there is no need for a context shift in grounded sentences, they provide a natural candidate for the collection of non-expansion sentences.

Now consider the closing-off of the minimal Kripkean fixed-point. The closing-off contains all sentences φ that are *classically valid* in the base model $\mathbb{M}(X)$ together with the extension of the minimal fixed-point. Let E be the extension of the truth predicate according to the minimal fixed-point, then the closing-off is given by the set $\{\varphi \mid \langle\mathbb{M}(X), \mathsf{E}\rangle \models \varphi\}$. As noted in section 3.2.4, since $\lambda \notin \mathsf{E}$, $\langle\mathbb{M}(X), \mathsf{E}\rangle \not\models \mathsf{Tr}(\lfloor\lambda\rfloor)$. Since the closing-off is fully classical $\langle\mathbb{M}(X), \mathsf{E}\rangle \models \neg\mathsf{Tr}(\lfloor\lambda\rfloor)$, and so by definition of λ, $\langle\mathbb{M}(X), \mathsf{E}\rangle \models \lambda$. Consequently, unlike the minimal fixed-point, the closing-off validates some expansion sentences. But this provides us with a criterion to identify expansion sentences: expansion sentences are the sentences that are validated by the closing-off but not contained in the minimal Kripkean fixed-point. This provides an answer to question (1), which sentences should be classified as expansion sentences. Note that this is a particularly elegant answer because the notion of expansion sentences is not presupposed from the outset, but rather drops out of the semantic theory.

To further refine this, start with an initial context c_0 and a contextualized version of the liar sentence λ_0 along with a contextualized truth predicate Tr_0. Construct the minimal Kripkean fixed-point over a model with the universal ´domain' and use it as the absolutist part of the semantics. Then, construct the closing-off over an arithmetical *set-sized* base structure $\mathcal{M}_0$, and define the relativist part of the semantics as the (relativistic) closing-off minus the sentences in the absolutist part (the minimal fixed-point constructed over the universal ´domain'). This allows us to define a bicontextualist notion of logical consequence for the initial context, where non-expansion sentences are interpreted by the absolutist part and expansion sentences by the relativist part of the semantics.

As for the context shift, consider c_1 and construct the minimal Kripkean fixed point for c_1, and take it as the absolutist part of the semantics. Similarly for the relativistic part: start from an acceptable, set-sized structure $\mathcal{M}_1$ that expands $\mathcal{M}_0$ (in a sense yet to be made precise), and construct the (relativistic) closing-off. Again, take the sentences that are in the relativistic $\mathcal{M}_1$-closing-off *and not* in the absolutist semantics of c_1 as the relativistic part of the semantics.

As in the initial context, we can define a notion of logical consequence that picks out, for each sentence, the *right* interpretation.

The described procedure yields the following picture: we get, on the relativistic side, a sequence of increasing, but set-sized models

$$\mathcal{M}_0 \lhd \mathcal{M}_1 \lhd ... \lhd \mathcal{M}_\alpha \lhd ...$$

each of which comes with a (relativistic) closing-off that we use to construct the relativist parts of the semantics. I will denote this by Rel_α. On the absolutist side, we get a sequence of minimal Kripkean fixed points, constructed for each language $\mathcal{L}_\alpha$, constructed over an absolutist-friendly, acceptable structure. This yields the absolutist parts of the semantics, denoted Abs_α. Finally, logical consequence can be defined as a relation between a set $\Gamma \subseteq \mathsf{Abs}_\alpha \cup \mathsf{Rel}_\alpha$ and a sentence φ as preservation of *absolute or relative truth*.

As mentioned at the end of section 3.1.3, my working hypothesis is to adopt the plural interpretation of higher-order expressions, as this fits naturally with the need for an absolutist-friendly interpretation of higher-order languages. Section 3.1.3 also provides a more detailed introduction to plural logic and discusses the ontological innocence of plural reference. For now, suffice it to say that the plural interpretation is used to interpret *n*th-order variables in terms of superpluralities, superduperpluralities, and so on. This allows me to have ontologically innocent higher-order quantification without reifying collection-like entities. Although my approach draws on the plural interpretation introduced by Boolos, it departs from his somewhat skeptical attitude toward higher-order pluralities.[15] In this work, higher-order pluralities are treated as legitimate and ontologically innocent referents of higher-order expressions. A more detailed discussion of an absolutist-friendly interpretation of higher-order expressions will be given in section 6.1.3.

4.3 TECHNICAL PRESENTATION OF HIGHER-ORDER BICONTEXTUALISM

In this section, I give the technical presentation of higher-order bicontextualism, extending the absolutist and relativist semantics introduced in Rossi [144]. To achieve this, I build upon the RU approach discussed in section 3.4. However, since only the core structure of the RU approach was introduced in section 3.4, several modifications are required to fully accommodate truth-theoretic reasoning. Specifically, this involves refining the definition of object languages in two key respects: first, by incorporating a sequence of context-relative truth predicates, Tr_α, and second, by ensuring that these languages have acceptable structures in the sense of section 3.2.1 (appropriately adjusted to higher-order languages and translated into the higher-order framework). As a result, the object languages developed here will more closely resemble those formulated for contextualism (section 3.2.4) rather than those originally defined within the RU approach (section 3.4).

After refining the definition of the object languages, I translate the acceptability requirements into the RU framework (section 4.3.1). This acceptability

15 See especially Boolos [16, 12].

requirements is crucial for the definition of a Kripke fixed-point in the RU framework. Next, I modify the definition of the satisfaction predicate to include additional clauses for the truth predicate (section 4.3.2). The revised satisfaction definition will serve as the foundation for the absolutist aspect of higher-order bicontextualist semantics. Following this, I develop the relativist component of the semantics (section 4.3.3) and, finally, introduce the anticipated bicontextualist account of logical consequence (section 4.3.4).

I begin with the object languages, focusing primarily on introducing, for each context c_α, a new unary predicate, Tr_α, which represents object-language truth of context c_α. Here is the definition followed by an explanatory remark.

Definition 4.1. Let $\mathcal{L}_\alpha := \mathcal{L}^1_\alpha, \mathcal{L}^2_\alpha, ..., \mathcal{L}^n_\alpha, ...$ be a sequence of higher-order languages such that $\mathcal{L}^n_\alpha$ is the n^{th}-order language of context α, and let

$$\mathfrak{L} := \bigcup_{\alpha \in Ord} \mathcal{L}_\alpha$$

be the collection of all these languages such that for each s and ordinals $\beta < \alpha$, the languages $\mathcal{L}^s_\alpha$, $\mathcal{L}^s_\beta$ satisfy the following requirements:

(a) $\mathcal{L}^s_\alpha$ contains $\in$ but has no function symbols.

(b) Each sequence of languages $\mathcal{L}_\alpha$ has a fresh unary predicate, Tr_α, aiming to express *truth-in-*c_α.

(c) $\mathcal{L}^s_\alpha$ and $\mathcal{L}^{s+1}_\alpha$ have the same terms (individual and relation variables and constants) up to type $(s-1)$, including Tr_α, but $\mathcal{L}^{s+1}_\alpha$ additionally has terms of type-s.

(d) For every s and every α, there is at least one acceptable $\mathcal{L}_\alpha$-structure $\mathcal{M}_\alpha$.

(e) For every acceptable structure $\mathcal{M}_\alpha$, and every $\mathcal{L}_\alpha$ expression e, there exists a closed term $\lfloor e \rfloor$ which denotes the code of e in $\mathcal{M}_\alpha$.

(f) For every acceptable structure $\mathcal{M}_\alpha$, and every open $\mathcal{L}_\alpha$-formula $\varphi(x)$, there exists a closed term t_φ s.t. $(t_\varphi)^{\mathcal{M}_\alpha} = (\lfloor \varphi(t_\varphi/x) \rfloor)^{\mathcal{M}_\alpha}$.

Remark 4.2.

- By (a), each language contains the membership relation $\in$, allowing it to develop mathematical notions in their set-theoretic representation. As in definition 3.31, restricting the languages to include only relation symbols (excluding function symbols) simplifies the subsequent definitions without any loss.

- By (b), each context α has its own truth predicate Tr_α, which will be used to express contextualized truth. The predicate Tr_α applies to expressions of any order within context α.

- By (c), any two languages $\mathcal{L}^s_\alpha \subseteq \mathcal{L}^{s+1}_\alpha$ within $\mathcal{L}_\alpha$ differ only in that $\mathcal{L}^{s+1}_\alpha$ extends $\mathcal{L}^s_\alpha$ by introducing new type-s terms.

- By (b) and (c), $\mathcal{L}_\alpha \subset \mathcal{L}_{\alpha+1}$, meaning that when shifting from c_α to $c_{\alpha+1}$, no expressions from c_α are removed. Instead, the new context $c_{\alpha+1}$ expands c_α by at least introducing a new truth predicate $\mathsf{Tr}_{\alpha+1}$.
- By (d), each context has at least one acceptable structure in the sense of section 3.2.1. This is a higher-order structure for c_α, denoted $\mathcal{M}^s_\alpha$, which interprets the natural numbers (including the natural ordering) within its first-order reduct. Further discussion on this can be found in section 4.3.1.
- By (e) and (f), every language has sufficient terms to allow for the coding of names and the formulation of self-referential sentences, such as the liar sentence λ.

4.3.1 *Acceptable Higher-Order RU Models*

The next step in the development of bicontextualism for higher-order object languages is to ensure that the higher-order RU models meet the acceptability criteria as defined by Moschovakis.[16] In general terms, the acceptability requirement asserts that a structure must be capable of interpreting $\langle \mathbb{N}, S \rangle$ such that an elementary coding scheme can be defined.

Ensuring the acceptability of higher-order RU models involves two steps: First, we need to extend the acceptability requirement, originally formulated for first-order languages in section 3.2.1, to higher-order languages. A crucial part of this process is to treat higher-order languages as multisorted languages, allowing us to generalize the acceptability condition so that mixed tuples can be coded. Second, we must translate the extended acceptability requirement into the formal language of the RU approach.

Let me begin by addressing the lifting of the acceptability requirement:

A coding function is a mapping from finite sequences of objects in the domain to a single object in the domain. In the context of higher-order languages, the term 'object' can mean different things, and the finite sequence can be *mixed*. In other words, a coding for an n^{th}-order model maps tuples of objects from orders up to order n to a single object of order n. This coding is an $(n+1)^{\text{th}}$-order object, defined in the $(n+1)^{\text{th}}$-order metalanguage as:

$$\lfloor \cdot \rfloor^{n+1} : \langle x^i_1, \ldots, x^j_m \rangle \mapsto x^n$$

for $i, j \in \{0, \ldots, n\}$. Importantly, the definition of $\lfloor \cdot \rfloor^{n+1}$ must be elementary, i.e., $(n+1)^{\text{th}}$-order definable.

The simplest case arises when we have a two-sorted language, as in Moschovakis [106, p. 23], where one sort is for natural numbers m, n, l, and the other for real numbers α, β, γ, interpreted as functions $\omega \mapsto \omega$. In this case, we can represent each natural number n as a constant function α, thereby allowing us to 'lift' natural numbers to the real number type. A tuple, whether mixed or not, can now be encoded by an object of the real number type. That is, the coding maps tuples to real numbers, such that $\langle \alpha, \beta \rangle \mapsto \gamma$, where α and β can be either real

16 See section 3.2.1.

numbers or 'lifted' natural numbers. This approach can be generalized to n-tuples using Cantor's pairing function. Following Moschovakis, an elementary coding schema is definable for a two-sorted language of this kind.

In essence, the approach here is a generalization of Moschovakis' ideas. However, rather than working with a two-sorted language, I use a multisorted language with n sorts, where each type can be appropriately lifted to higher types. The highest type contains an isomorphic copy of the natural numbers, which serves as the basis for the definition of an elementary coding schema.

If $n = 0$, I assume that the model contains a definable subset N^1 that is isomorphic to the natural numbers. This implies the existence of a distinguished constant $0^0 \in N^1$, which is interpreted as the number zero, and a *successor function* $S^1 : N^1 \to N^1$ that satisfies $S^1(x^0) \neq 0^0$ and is injective.[17] Therefore, the triple $(N^1, S^1, 0^0)$ forms a copy of the natural numbers at type-0.[18]

More generally, for each $n \geq 0$, I assume that there is a distinguished predicate $N^{n+1}(x^n)$ that defines a subset of the n^{th}-order domain, which we interpret as the copy of $\mathbb{N}$ at level n. Specifically, there must be a distinguished constant 0^n such that $N^{n+1}(0^n)$ holds. Intuitively, 0^n represents the 'lifted zero' at type n. Additionally, there exists an elementary function (on type n), S^{n+1}, the 'lifted successor function', which is injective and, for every $x^n \in N^{n+1}$, satisfies $S^{n+1}(x^n) \in N^{n+1}$. Furthermore, it is required that 0^n is not in the range of S^{n+1}.[19] These assumptions ensure that each type has an *internal copy* of $\mathbb{N}$.

Having a 'lifted' copy of $\mathbb{N}$ at level n, mixed tuples can now take the form $\langle x^i, y^j \rangle$, where $i, j \in \{0, \ldots, n\}$. A coding of $\langle x^i, y^j \rangle$ operates on the 'lifted' objects representing x^i and y^j at level n. Specifically, a coding is a mapping $\langle x^n, y^n \rangle \mapsto z^n$, where x^n and y^n are either objects of type n or represent objects of lower types within type n, i.e., they are 'lifted' objects. As in the case of a two-sorted language, n-tuples can be handled in the standard way using Cantor's pairing function. Moreover, such a function is elementarily definable in the $(n+1)^{\text{th}}$-order language.

The next step is to translate the acceptability requirement into the formal framework of the RU approach. The definition must ensure that the distinguished constants and predicates mimicking $\mathbb{N}$ *behave arithmetically*. To achieve this, I use the formula $\mathbb{N}(N^{n+1}, S^{n+1}, 0^n)$. A full formalization of this involves ensuring that N^{n+1} is Dedekind infinite, i.e., there exists an injection from N^{n+1} into itself, and that S^{n+1} is a function from N^{n+1} to N^{n+1} which does not have 0^n in its range.

Taken together, this yields the following: a type-n coding scheme is a mapping from sequences of type-n objects to a single type-n object such that the following

- $\mathsf{seq}^{n+1}(x^n)$, the set of codes of finite sequence;

17 From this, the binary relation $<^1 \subseteq N^1 \times N^1$, which satisfies the usual properties of the standard ordering on $\mathbb{N}$, is elementarily definable.

18 Note that, since I am working exclusively with relational languages, all these functions must be interpretable in the sense that a relation is definable, which is univocal to the right and can mimic the specific function. See section 3.2.1.

19 Again, the binary relation $<^{n+1}$ on N^{n+1}, satisfying the usual order properties of the natural ordering, this time on N^{n+1}, is definable.

- $\mathsf{lh}^{n+1}(x^n) = \begin{cases} 0^n, & \text{if } \neg\mathsf{seq}^{n+1}(x^n) \\ m^n, & \text{if } \mathsf{seq}^{n+1}(x^n) \wedge x = \lfloor x_1^n, ..., x_m^n \rfloor^{n+1} \end{cases}$
- $\mathsf{q}^{n+1}(x^n, k^n) = \begin{cases} x_i^n, & \text{if } x^n = \lfloor x_1^n, ..., x_i^n, ...x_m^n \rfloor^{n+1} \wedge 1 \leq i \leq m \\ 0^n, & \text{otherwise.} \end{cases}$

are definable. A RU model $\mathbb{M}(X^{n+1})$ is *acceptable* if it admits an elementary type-n coding scheme.[20]

I will now give the translation of the acceptability requirement into the RU approach. The definition should be such that we can add it as a conjunct to the model predicate in order to ensure that the model is acceptable. An explanatory remark follows afterwards:

Definition 4.3. For every unary type-$(n+1)$-variable X^{n+1} such that $\mathbb{M}(X^{n+1})$, we define the acceptability condition, in symbols $\mathbb{A}(X^{n+1})$, as:

$$
\begin{aligned}
&\mathbb{A}(X^{n+1}) :\leftrightarrow \exists N^{n+1}\exists S^{n+1}\exists 0^n\Big(\\
&\quad \mathbb{N}(N^{n+1}, S^{n+1}, 0^n) \wedge N^{n+1} \preceq Y^{n+1} \wedge S^{n+1} \preceq Y^{n+1}\Big)\wedge \\
&\exists \lfloor\cdot\rfloor^{n+1}\Big[\lfloor\cdot\rfloor^{n+1} \preceq Y^{n+1} \wedge \forall x_1^n, ..., x_m^n, x^n\Big(Y^{n+1}(x_1^n) \wedge ... \wedge Y^{n+1}(x_m^n)\wedge \\
&\quad N^{n+1}(x^n) \wedge func(\lfloor\cdot\rfloor^{n+1}) \wedge inj(\lfloor\cdot\rfloor^{n+1}) \wedge \lfloor x_1^n, ..., x_m^n\rfloor^{n+1} = x^n\Big)\Big]\wedge \\
&\exists \mathsf{sq}^{n+1}\Big[\mathsf{sq}^{n+1} \preceq Y^{n+1} \wedge \forall x^n\Big(\mathsf{sq}^{n+1}(x^n) \leftrightarrow \exists x_1^n, ..., x_m^n(\lfloor x_1^n, ..., x_m^n\rfloor^{n+1} = x^n)\Big)\Big]\wedge \\
&\exists \mathsf{lh}^{n+1}\Big[\mathsf{lh}^{n+1} \preceq Y^{n+1} \wedge \forall x^n\Big(\big(\neg\mathsf{sq}^{n+1}(x^n) \rightarrow \mathsf{lh}^{n+1}(x^n) = 0^n\big)\vee \\
&\quad \big(\mathsf{sq}^{n+1}(x^n) \rightarrow \exists m^n, x_1^n, ..., x_m^n(N^{n+1}(m^n) \wedge Y^{n+1}(x_1^n) \wedge ... \wedge Y^{n+1}(x_m^n)\wedge \\
&\quad \lfloor x_1^n, ..., x_m^n\rfloor^{n+1} = x^n \wedge \mathsf{lh}^{n+1}(x^n) = m^n)\big)\Big)\Big]\wedge \\
&\exists \mathsf{q}^{n+1}\Big[\mathsf{q}^{n+1} \preceq Y^{n+1} \wedge \forall x^n, i^n\Big(N^{n+1}(i^n) \wedge \mathsf{sq}(x^n) \rightarrow \\
&\quad ((\mathsf{lh}^{n+1}(x^n) < i^n \rightarrow \mathsf{q}^{n+1}(x^n, i^n) = 0^n) \vee (\mathsf{lh}^{n+1}(x^n) \geq i^n \rightarrow \\
&\quad \exists x_1^n, ..., x_i^n, ..., x_m^n(\lfloor x_1^n, ..., x_i^n, ..., x_m^n\rfloor^{n+1} = x^n \wedge \mathsf{q}^{n+1}(x^n, i^n) = x_i^n))))\Big]
\end{aligned}
$$

To say that X^{n+1} *is an acceptable RU structure* is short for $\mathbb{M}(X^{n+1}) \wedge \mathbb{A}(X^{n+1})$.

20 Note that there is a subtle shift in the elementarity requirement: in the first-order case, the requirement is that a function $\lfloor\cdot\rfloor$ is definable within the first-order language. However, in a type-theoretic setting, a function symbol is itself an object of type 1—that is, an expression in a second-order language. More generally, a function $\lfloor\cdot\rfloor^{n+1}$ that maps tuples of type-n elements to a single type-n element is an object of type $(n+1)$, meaning it can only be part of the language of order $(n+1)$. While it would be possible to simulate the transition from first- to second-order logic in the case of a type-n and a type-$(n+1)$ language—by introducing type-$(n+1)$ constants into the type-n language while restricting quantification to type-n variables—such an approach would introduce unnecessary complexity in the definition of the languages. Instead, the most natural formulation of the elementarity requirement is that the coding function is definable in $\mathcal{L}^{n+1}$.

Remark 4.4.

- First, it is certainly unusual to quantify over expressions that resemble names rather than variables, as I have done multiple times above ($\exists N^{n+1}$, $\exists S^{n+1}$, $\exists 0^{n}$, $\exists \lfloor \cdot \rfloor^{n+1}$, $\exists \mathsf{sq}^{n+1}$, $\exists \mathsf{lh}^{n+1}$, and $\exists \mathsf{q}^{n+1}$). However, this is simply to improve readability, and one should always remember that all these expressions are formally treated as variables.
- The first part of the definition ensures that the objects representing the natural numbers lifted to type-n exist and that they form an arithmetical structure. As anticipated, this is accomplished by $\mathbb{N}(N^{n+1}, S^{n+1}, 0^{n})$, which I assume to be given.[21] In the definition, Y^{n+1} represents the domain of the model X^{n+1}.
- The next line introduces the coding function $\lfloor \cdot \rfloor^{n+1}$, which operates on objects of order n (where all objects are either of order n or are 'lifted' appropriately) and maps them to a single object of order n that satisfies the 'lifted natural number predicate'. Additionally, $func(\lfloor \cdot \rfloor)$ and $inj(\lfloor \cdot \rfloor)$ indicate that $\lfloor \cdot \rfloor^{n+1}$ is functional and injective.
- The following three lines internalize the sequence predicate sq^{n+1}, as well as the functions for length lh^{n+1} and projection q^{n+1}. Since these definitions rely on the definition of $\lfloor \cdot \rfloor^{n+1}$, they impose additional requirements on $\lfloor \cdot \rfloor^{n+1}$, similar to the standard model-theoretic definition of an acceptable structure, as given in Moschovakis [106, ch. 5].

4.3.2 *A Kripkean Truth Predicate for Unrestricted Higher-Order Languages*

Having generalized the notion of an acceptable structure to capture RU models, I will now use the strong Kleene satisfaction predicate (definition 3.44) to define a *Kripke Kleene satisfaction predicate (strong version)* (KKSat), which generalizes the strong Kleene satisfaction predicate SKSat to capture the notion of object language truth. Essentially, the definition introduces clauses for the truth predicate.

As pointed out in section 3.2.3, the main idea of Kripke's seminal paper [79] is to build up the extension of the truth predicate as the fixed-point of a sequence $E_0, E_1, ..., E_\alpha, ...$ of (codes of) sentences of an arithmetical base language with a truth predicate. The extension is built model-theoretically by first considering all atomic sentences of the truth-free fragment of the base language and evaluating them over an acceptable base structure $\mathcal{M}$. The first step yields a set E_1 containing all and only (codes of) atomic (truth-free) sentences that are true in $\mathcal{M}$, such as $n = n$, and (codes of) negations of atomic (truth-free) sentences that are false in $\mathcal{M}$, such as $\neg(n = m)$. Then, we go from E_1 to E_2 by applying the closure operations according to strong Kleene satisfaction *and truth-application*. So, if φ and ψ are in E_1, then $\neg\neg\varphi$, $\varphi \wedge \psi$, $\mathsf{Tr}(\lfloor \varphi \rfloor)$, and so on are in E_2.

To make the Kripke construction compatible with absolutism, I need to translate it into the higher-order framework. Since the primary goal is to

21 A formal definition can be found in [18, Ch. 10].

recreate Rossi's version of Glanzberg's contextualism, I also need to expand the definition to allow for non-minimal fixed points. This is done by allowing an additional higher-order parameter, which can be used as the higher-order analogue of the set that serves as the starting point for constructing non-minimal fixed points in the standard setting (see section 3.2.4). Moreover, since I work with a sequence of contextualized higher-order object languages $\mathcal{L}^n_\alpha$, and since contextualism allows for constructing a sequence of non-minimal fixed points, I present the definition here for a fixed contextualized language $\mathcal{L}^n_\alpha$. Let me first provide the definition and offer an explanatory remark afterwards:

Definition 4.5. Let $\mathcal{L}^n_\alpha$ be a type-n language for c_α, let $i, j \in \{0, \ldots, n\}$, let x^n be a type-n variable, and X^{n+1} and Y^{n+1} be type-$(n+1)$ variables. The $\mathcal{L}^{n+1}_\alpha$-formula "the RU-model X^{n+1} with RU-domain Y^{n+1} Kripke-Kleene-satisfies x^n relative to the accepted Z^{n+1}", in symbols $\mathsf{KKSat}^n_\alpha(x^n, X^{n+1}, Y^{n+1}, Z^{n+1})$, is defined as: $\mathsf{KKSat}^n_\alpha(x^n, X^{n+1}, Y^{n+1}, Z^{n+1})$:iff

1. $\mathbb{M}(X^{n+1})$ and $\mathbb{A}(X^{n+1})$ and $\mathbb{D}(X^{n+1}, Y^{n+1})$ and $x^n \in \mathsf{Sent}_{\mathcal{L}^n_\alpha}$ and
2. $Z^{n+1}(x^n)$, or
3. x^n is $z^s(t^i_1, ..., t^j_m)$ and $X^{s+1}(\langle z^s, \langle t^i_1, ..., t^j_m\rangle\rangle)$, or
4. x^n is $\dot\neg z^s(t^i_1, ..., t^j_m)$ and $\neg X^{s+1}(\langle z^s, \langle t^i_1, ..., t^j_m\rangle\rangle)$, or
5. x^n is $\dot\neg\dot\neg y^n$ for $y^n \in \mathsf{Sent}_{\mathcal{L}^n_\alpha}$ and $\mathsf{KKSat}^n_\alpha(y^n, X^{n+1}, Y^{n+1}, Z^{n+1})$, or
6. x^n is $y^n \dot\wedge z^n$ for $y^n \in \mathsf{Sent}_{\mathcal{L}^n_\alpha}$ and $z^n \in \mathsf{Sent}_{\mathcal{L}^n_\alpha}$ and $\mathsf{KKSat}^n_\alpha(y^n, X^{n+1}, Y^{n+1})$ and $\mathsf{KKSat}^n_\alpha(z^n, X^{n+1}, Y^{n+1}, Z^{n+1})$, or
7. x^n is $\dot\neg(y^n \dot\wedge z^n)$ for $y^n \in \mathsf{Sent}_{\mathcal{L}^n_\alpha}$ and $z^n \in \mathsf{Sent}_{\mathcal{L}^n_\alpha}$ and $\mathsf{KKSat}^n_\alpha(\dot\neg y^n, X^{n+1}, Y^{n+1}, Z^{n+1})$ or $\mathsf{KKSat}^n_\alpha(\dot\neg z^n, X^{n+1}, Y^{n+1}, Z^{n+1})$, or
8. x^n is $\dot\forall z^j y^n(z^j)$ for $y^n(z^j) \in \mathsf{For}_{\mathcal{L}^n_\alpha}$, and for every W^{n+1} s.t. $\mathbb{V}(z^j, W^{n+1}, X^{n+1})$, $\mathsf{KKSat}^n_\alpha(y^n(z^j), W^{n+1}, Y^{n+1}, Z^{n+1})$, or
9. x^n is $\dot\neg\dot\forall z^j y^n(z^j)$ for $y^n(z^j) \in \mathsf{For}_{\mathcal{L}^n_\alpha}$, and for some W^{n+1} s.t. $\mathbb{V}(z^j, W^{n+1}, X^{n+1})$, $\mathsf{KKSat}^n_\alpha(y^n(z^j), W^{n+1}, Y^{n+1}, Z^{n+1})$, or
10. x^n is $\mathsf{Tr}_\alpha(y^n)$ and y^n is $\lfloor z^n \rfloor^{n+1}$ for $z^n \in \mathsf{Sent}_{\mathcal{L}^n_\alpha}$ and $\mathsf{KKSat}^n_\alpha(z^n, X^{n+1}, Y^{n+1}, Z^{n+1})$, or
11. x^n is $\dot\neg\mathsf{Tr}_\alpha(y^n)$ and either does y^n not code an $\mathcal{L}^n_\alpha$-sentence, or y^n is $\lfloor z^n \rfloor^{n+1}$ for $z^n \in \mathsf{Sent}_{\mathcal{L}^n_\alpha}$ and $\mathsf{KKSat}^n_\alpha(\dot\neg z^n, X^{n+1}, Y^{n+1}, Z^{n+1})$.

Remark 4.6.

- Unlike the general definition given in section 3.2.2 (or the Kripke version presented in section 3.2.3), where fixed points are certain stages in an ordinal-indexed sequence of increasing sets, definition 4.5 defines the fixed point directly without the sequence of growing initial segments. The two approaches are equivalent (see [57, Ch. 15]), and adopting the latter approach simplifies the presentation here without altering the results.

- The first clause of the definition of $\mathsf{KKSat}^n_\alpha(x^n, X^{n+1}, Y^{n+1}, Z^{n+1})$ ensures that X^{n+1} is an acceptable RU model with domain Y^{n+1}, and that the definition operates on (names of) sentences in the object language $\mathcal{L}^n_\alpha$.
- The second clause, $Z^{n+1}(x^n)$, will be used to define non-minimal fixed points. Just as in the standard definition of contextualism (section 3.2.4), we can use the (possibly non-minimal) fixed point of context c_α to define the extension of the truth predicate of the subsequent fixed point $c_{\alpha+1}$ as the non-minimal fixed point starting from $\mathsf{KKSat}^n_\alpha(x^n, X^{n+1}, Y^{n+1}, Z^{n+1})$. If $Z^{n+1} = \emptyset$, the definition of $\mathsf{KKSat}^n_\alpha(x^n, X^{n+1}, Y^{n+1}, \emptyset)$ yields the minimal fixed point of c_α.
- According to definition 4.5, the extension of KKSat^n_α contains all true $\mathcal{L}^n_\alpha$-sentences of the RU model X^{n+1} and RU domain Y^{n+1}. Taking X^{n+1} to be an absolutist 'domain', KKSat^n_α directly generates the extension of the (possibly non-minimal) Kripkean fixed point for a type-n language with quantifiers ranging over an unrestricted 'domain'.

The $\mathcal{L}^{n+1}_\alpha$-formula KKSat^n_α defines the notion of a Kripkean fixed point, constructed over an absolutist-friendly base model with unrestricted 'domain' that satisfies the acceptability requirement. In the next step, I use this to define the notion of *absolutely general type-n truths of* c_α. Recall from section 4.2 that the absolutist part of the semantics is given by the minimal Kripkean fixed point, as it contains only grounded sentences. As described there, I will use this in the following step to define the relativist part of the semantics. Here is the definition of the absolutist part:

Definition 4.7. For every $\mathcal{L}^n_\alpha$, the set of *absolutely general truths* of $\mathcal{L}^n_\alpha$ is defined as:

$$\mathsf{Abs}^n_\alpha := \{\varphi \in \mathcal{L}^n_\alpha \mid \mathsf{KKSat}^n_\alpha(\lfloor\varphi\rfloor^{n+1}, X^{n+1}, Y^{n+1}, \emptyset)\}.$$

I take as the absolutely general $\mathcal{L}^n_\alpha$-truths of context c_α the sentences in the minimal fixed point of the context c_α, as constructed over an absolutist-friendly RU model X^{n+1} with an unrestricted RU 'domain' Y^{n+1}. Note that, by this definition, each context has its own absolutist semantics. After the definition of the relativistic semantics in the next section, this allows us to define, for each context c_α, a bicontextualist notion of logical consequence, in which each context has the ability to evaluate some sentences over the unrestricted 'domain', and others only over a restricted domain.

4.3.3 *Relativist Semantics*

Let me now define the relativist part of the semantics. Recall from section 3.1.2 that a model (in the sense of model theory) for an n^{th}-order object language is denoted by $\mathcal{M}^n$. To begin, we must recreate the sequence of contextualized (relativistic) higher-order closing-offs, one for each context c_α. To align with the principles of Rossi's simplified version of Parsons-Glanzberg contextualism, it is essential to work with a sequence of increasing, set-sized higher-order

models. These models build upon one another such that, for ordinals α, β with $\beta < \alpha$, the model $\mathcal{M}^n_\beta$ is embedded into $\mathcal{M}^n_\alpha$ (i.e., $\mathcal{M}^n_\beta \lhd \mathcal{M}^n_\alpha$ in a sense yet to be made precise). Ultimately, we want $\mathcal{M}^n_\alpha$ to agree with all previous models on their fixed points. Let me first translate the closing-off into the higher-order framework and explain it afterwards.

Definition 4.8. Let $\mathcal{L}^n_0$ be a type-n language of the initial context c_0, $\mathcal{M}^n_0$ be an acceptable model of the Tr_0-free fragment of $\mathcal{L}^n_0$, and M^n_0 its n^{th}-order domain. The relativistic closing-off of $\mathcal{L}^n_0$, relative to $\mathcal{M}^n_0$, is defined as:

$$\mathsf{ClOff}^n_0 := \{\varphi \in \mathcal{L}^n_0 \mid \langle \mathcal{M}^n_0, \{x^n \in M^n_0 \mid \mathsf{KKSat}^n_0(x^n, \mathcal{M}^n_0, M^n_0, \emptyset)\}\rangle \models \varphi\}.$$

More generally, for $\alpha > 0$, $\beta < \alpha$, let $\mathcal{M}^n_\alpha$ be a model of the truth-free fragment of $\mathcal{L}_\alpha$ such that

$$\mathcal{M}^n_\beta \lhd_{\bigcup_{\beta<\alpha} \mathsf{ClOff}^n_\beta} \mathcal{M}^n_\alpha,$$

i.e., $\mathcal{M}^n_\alpha$ and all the $\mathcal{M}^n_\beta$'s are elementary equivalent with respect to the sentences in $\bigcup_{\beta<\alpha} \mathsf{ClOff}^n_\beta$. Then, the c_α closing off is defined as:

$$\mathsf{ClOff}^n_\alpha := \{\varphi \in \mathcal{L}^n_\alpha \mid \langle \mathcal{M}^n_\alpha, \{x^n \in M^n_\alpha \mid \mathsf{KKSat}^n_\alpha(x^n, \mathcal{M}^n_\alpha, M^n_\alpha, \bigcup_{\beta<\alpha} \mathsf{ClOff}^n_\beta)\}\rangle \models \varphi\}.$$

Remark 4.9.

- Let's look at the definition of ClOff^n_0 first: Start from a *set-sized* acceptable base structure of the truth-free fragment of $\mathcal{L}^n_0$, $\mathcal{M}^n_0$, with n^{th}-order domain M^n_0, and construct $\mathsf{KKSat}^n_0(x^n, \mathcal{M}^n_0, M^n_0, \emptyset)$, i.e., the least Kripkean fixed point. The relativistic (because built over a set-sized domain) closing-off is defined as the set of *classically valid sentences* of the structure $\mathcal{M}^n_0$, together with KKSat^n_0 as the interpretation of the c_0 truth predicate Tr_0.

- In the next step, I take a set-sized, acceptable c_1-structure $\mathcal{M}^n_1$ of the Tr_1-free fragment of $\mathcal{L}^n_1$, and use it to construct a new (but this time *non-minimal*) fixed-point. Since we want this fixed-point to interpret sentences containing Tr_0 as well, we construct this fixed point *over* the c_0 closing-off. In the definition of KKSat^n_1 (definition 4.5), we substitute ClOff^n_1 for the $(n+1)^{\text{th}}$-order variable Z^{n+1} from clause 2. The closing-off of this definition yields again a set of classically valid $\mathcal{L}^n_1$-sentences, where KKSat^n_1 is used to interpret Tr_1.

- This process can now be iterated to generate a sequence of increasing closing-offs, ClOff^n_α, one for each context c_α. Note that all these closing-offs are constructed over set-sized structures, and I will use them in the next step to define the relativistic part of the semantics.

- The requirement that $\mathcal{M}^n_\alpha$ and all the $\mathcal{M}^n_\beta$'s are elementary equivalent with respect to the sentences in $\bigcup_{\beta<\alpha} \mathsf{ClOff}^n_\beta$ is to make sure that the models agree on all previously constructed fixed points. This makes the sequence of closing-offs consistent in the sense that for no ordinals α, β and $\varphi \in \mathcal{L}^n_\beta$, $\langle \mathcal{M}^n_\beta, \mathsf{KKSat}^n_\beta\rangle \models \mathsf{Tr}_\beta(\lfloor\varphi\rfloor)$ but $\langle \mathcal{M}^n_\alpha, \mathsf{KKSat}^n_\alpha\rangle \not\models \mathsf{Tr}_\beta(\lfloor\varphi\rfloor)$ (the second Tr_β is not a typo). In other words, if φ is false in c_β, it must not be possible to say in c_α that φ is *true-in-*c_β.

- To see that $\mathcal{M}^n_\alpha$ and $\mathcal{M}^n_\beta$ shall only agree on $\bigcup_{\beta<\alpha} \mathsf{ClOff}^n_\beta$ and not on all $\mathcal{L}^n_\beta$-sentences, assume that $\mathcal{M}^n_\alpha$ has an inaccessible cardinal, but in $\mathcal{M}^n_\beta$, all sets are smaller in size. The formalization of 'there is an inaccessible cardinal', however, is identical in $\mathcal{L}^n_\alpha$ and $\mathcal{L}^n_\beta$ since the truth predicate is not required (in fact, nothing beyond $\in$ is required, which is part of all languages). Abbreviate this sentence by φ. Consequently, $\mathcal{M}^n_\beta \not\models \varphi$ but $\mathcal{M}^n_\alpha \models \varphi$, and so it would be problematic to require that $\mathcal{M}^n_\beta$ and $\mathcal{M}^n_\alpha$ agree on all $\mathcal{L}^n_\beta$-sentences.

Despite the fact that the integration in the RU framework makes some version of absolutist-friendly contextualist reasoning already possible, this is not the main aspect of definition 4.8. The sequence of closing-offs allows us to recreate Parsons-Glanzberg contextualism. The combination with absolutism comes only in the following section.

Let me now show that higher-order bicontextualism handles the liar sentences exactly as Parsons-Glanzberg contextualism, and hence as in Rossi's original paper [144]. Consider the c_0 liar λ_0. By construction, λ_0 is not in the minimal fixed point, i.e., KKSat^n_0 does not hold of λ_0. Consequently, as in section 3.2.4,

$$\langle \mathcal{M}^n_0, \{x^n \in M_0 \mid \mathsf{KKSat}^n_0(x^n, \mathcal{M}^n_0, M^n_0, \emptyset)\}\rangle \not\models \mathsf{Tr}_0(\lfloor \lambda_0 \rfloor).$$

But since closing-off models are fully classical, we get

$$\langle \mathcal{M}^n_0, \{x^n \in M_0 \mid \mathsf{KKSat}^n_0(x^n, \mathcal{M}^n_0, M^n_0, \emptyset)\}\rangle \models \neg\mathsf{Tr}_0(\lfloor \lambda_0 \rfloor),$$

and by the definition of λ_0, we get

$$\langle \mathcal{M}^n_0, \{x^n \in M_0 \mid \mathsf{KKSat}^n_0(x^n, \mathcal{M}^n_0, M^n_0, \emptyset)\}\rangle \models \lambda_0.$$

From the subsequent context, however, when we construct the non-minimal c_1 fixed point together with the corresponding closing-off, c_1 makes the claim 'λ_0 does not express a true proposition in c_0' true, i.e.,

$$\langle \mathcal{M}^n_1, \{x^n \in M_1 \mid \mathsf{KKSat}^n_1(x^n, \mathcal{M}^n_1, M^n_1, \mathsf{ClOff}^n_0)\}\rangle \models \mathsf{Tr}_1(\lfloor \lambda_0 \rfloor).$$

However, now from the c_1 point of view, there is a new c_1 liar λ_1 which behaves analogously, i.e.,

$$\langle \mathcal{M}^n_1, \{x^n \in M_1 \mid \mathsf{KKSat}^n_1(x^n, \mathcal{M}^n_1, M^n_1, \mathsf{ClOff}^n_0)\}\rangle \models \lfloor \lambda_1 \rfloor,$$

but in c_2 we get

$$\langle \mathcal{M}^n_2, \{x^n \in M_2 \mid \mathsf{KKSat}^n_2(x^n, \mathcal{M}^n_2, M^n_2, \mathsf{ClOff}^n_0 \cup \mathsf{ClOff}^n_1)\}\rangle \models \mathsf{Tr}_2(\lfloor \lambda_1 \rfloor),$$

and so on.

With the higher-order version of the closing-off, we can now define the relativist part of the semantics. As previously discussed in section 4.2, relativistic truths are those statements that are *not* part of the minimal fixed point but are contained in the closing-off. Recall that we have defined the absolutist truths Abs^n_α to be given by the minimal absolutist-friendly Kripkean fixed point (definition 4.7). Let me now define the relativistic truths and add an explanatory remark afterwards:

Definition 4.10. For every $\mathcal{L}^n_\alpha$, the set of *relatively general truths* of $\mathcal{L}^n_\alpha$ is defined as:

$$\mathsf{Rel}^n_\alpha := \mathsf{ClOff}^n_\alpha \setminus (\mathsf{Abs}^n_\alpha \cup \{\varphi \in \mathcal{L}^n_\alpha \mid \neg\varphi \in \mathsf{Abs}^n_\alpha\}).$$

Remark 4.11.

- The basis for the relativistic n^{th}-order truths of context α is the closing-off of an acceptable (set-sized) n^{th}-order c_α-structure $\mathcal{M}^n_\alpha$. However, we must filter out some sentences. First, all sentences that are also in Abs^n_α. For instance, both the set-sized *and* the absolutist-friendly model will make the claim $\forall x\ x = x$ true. However, since this claim is unproblematic in the sense discussed in chapter 2, the bicontextualist wants to have it evaluated over the all-inclusive 'domain'. Consequently, in order to construct here only the relativistic half of the semantics, we remove it (and with it, all sentences that behave similarly) from Rel^n_α.
- Moreover, we have to remove all sentences that are *absolutely general falsities* (i.e., the complement of Abs^n_α). Consider, again, a situation where two models might disagree on cardinalities. If there are inaccessible cardinals, the claim 'there exists no inaccessible cardinal' will certainly come out false over the all-inclusive domain. However, if there are only smaller sets in $\mathcal{M}^n_\alpha$, the claim will come out true there. But in this case, we should evaluate the sentence over the all-inclusive domain.[22]

In the next section, I will give the full picture of bicontextualism and define a bicontextual notion of logical consequence.

4.3.4 *A Bicontextualist Notion of Logical Consequence*

I will now define a notion of logical consequence in the sense of higher-order bicontextualism. The main idea is to define higher-order bicontextualist logical consequence as preservation of absolute n^{th}-order truths or relative n^{th}-order truth. This way, for an argument from Γ to φ to be *higher-order bicontextually valid*, the interpretation has to respect the expansion (or non-expansion) character of the sentences. In other words, in context α, the semantics should take all non-expansion sentences from Abs^n_α, and all expansion sentences from Rel^n_α:

Definition 4.12. For every $\mathcal{L}^n_\alpha$, and $\Delta \subseteq \mathsf{Sent}^n_\alpha$ and $\varphi \in \mathsf{Sent}^n_\alpha$, the argument from Δ to φ is *higher-order bicontextually valid*, in symbols $\Delta \models^n_\alpha \varphi$ iff: if all the sentences in Δ are in $\mathsf{Abs}^n_\alpha \cup \mathsf{Rel}^n_\alpha$, then so is φ.

The higher-order bicontextualist notion of logical consequence has several desirable features. Since bicontextualism is built on a succession of closed-off models, each context validates a higher-order version of KF, suitably adjusted to higher-order languages and for the respective context:[23]

22 Note that large cardinal claims are treated differently in set-theoretic bicontextualism (see chapter 5).

23 For the notational conventions concerning the dot-notation, $\mathsf{val}(x)$, and $x^0(t^s(v^s))$, see section 3.2.5.

Definition 4.13. Let $s, n, i, j \in \mathbb{N}$ such that $i, j < s \leq n$. The theory KF^n_α is given by all the axioms of *Peano Arithmetic* with induction in the language $\mathcal{L}^n_\alpha$, together with the following axioms:

$\mathsf{KF1}^n_\alpha \quad \forall x^s \forall y^i_1 ... \forall y^j_m (\mathsf{Tr}_\alpha(\dot{x}^s(y^i_1, ..., y^j_m)) \leftrightarrow \mathsf{val}(x^s)(\mathsf{val}(y^i_i), ..., \mathsf{val}(y^j_m))$

$\mathsf{KF2}^n_\alpha \quad \forall x^s \forall y^i_1 ... \forall y^j_m (\mathsf{Tr}_\alpha(\dot{x}^s(y^i_1, ..., y^j_m)) \leftrightarrow \neg\mathsf{val}(x^s)(\mathsf{val}(y^i_i), ..., \mathsf{val}(y^j_m))$

$\mathsf{KF3}^n_\alpha \quad \forall x^n (\mathsf{Sent}_{\mathcal{L}^n_\alpha}(x^n) \rightarrow (\mathsf{Tr}_\alpha(\dot{\neg}\dot{\neg}x^n) \leftrightarrow \mathsf{Tr}_\alpha(x^n)))$

$\mathsf{KF4}^n_\alpha \quad \forall x^n \forall y^n (\mathsf{Sent}_{\mathcal{L}^n_\alpha}(x^n \dot{\wedge} y^n) \rightarrow (\mathsf{Tr}_\alpha(x^n \dot{\wedge} y^n) \leftrightarrow \mathsf{Tr}_\alpha(x^n) \wedge \mathsf{Tr}_\alpha(y^n)))$

$\mathsf{KF5}^n_\alpha \quad \forall x^n \forall y^n (\mathsf{Sent}_{\mathcal{L}^n_\alpha}(x^n \dot{\wedge} y^n) \rightarrow (\mathsf{Tr}_\alpha(\dot{\neg}(x^n \dot{\wedge} y^n)) \leftrightarrow \mathsf{Tr}_\alpha(\dot{\neg}x^n) \vee \mathsf{Tr}_\alpha(\dot{\neg}y^n)))$

$\mathsf{KF6}^n_\alpha \quad \forall v^s \forall x^n (\mathsf{Sent}_{\mathcal{L}^n_\alpha}(\dot{\forall} v^s x^n) \rightarrow (\mathsf{Tr}_\alpha(\dot{\forall} v^s x^n) \leftrightarrow \forall t^s \mathsf{Tr}_\alpha(x^n(t^s/v^s))))$

$\mathsf{KF7}^n_\alpha \quad \forall v^s \forall x^n (\mathsf{Sent}_{\mathcal{L}^n_\alpha}(\dot{\forall} v^s x^n) \rightarrow (\mathsf{Tr}_\alpha(\dot{\neg}\dot{\forall} v^s x^n) \leftrightarrow \exists t^s \mathsf{Tr}_\alpha(\dot{\neg}x^n(t^s/v^s))))$

$\mathsf{KF8}^n_\alpha \quad \forall t^n (\mathsf{Tr}_\alpha(\dot{\mathsf{Tr}}_\alpha(t^n)) \leftrightarrow \mathsf{Tr}_\alpha(\mathsf{val}(t^n)))$

$\mathsf{KF9}^n_\alpha \quad \forall t^n (\mathsf{Tr}_\alpha(\dot{\neg}\dot{\mathsf{Tr}}_\alpha(t^n)) \leftrightarrow (\mathsf{Tr}_\alpha(\dot{\neg}\mathsf{val}(t^n)) \vee \neg\mathsf{Sent}_{\mathcal{L}^n_\alpha}(\mathsf{val}(t^n))))$

Theorem 4.14. Higher-order bicontextualism validates KF^n_α, i.e.,

$$\models^n_\alpha \mathsf{KF}^n_\alpha.$$

Proof. We show that for each $\varphi \in \mathsf{KF}^n_\alpha$, $\lfloor \varphi \rfloor^{n+1} \in \mathsf{Abs}^n_\alpha \cup \mathsf{Rel}^n_\alpha$. The claim follows by induction on the complexity of formulae. For the base case, assume that $x^s(y^i_1, ..., y^j_m)$ holds in c_α, i.e., one of the following obtains:

$$\mathsf{KKSat}^n_\alpha(\lfloor x^s(y^i_1, ..., y^j_m) \rfloor^{n+1}, X^{n+1}, Y^{n+1}, \varnothing) \tag{4.1}$$

$$\mathsf{KKSat}^n_\alpha(\lfloor x^s(y^i_1, ..., y^j_m) \rfloor^{n+1}, \mathcal{M}^n_\alpha, M^n_\alpha, \bigcup_{\beta<\alpha} \mathsf{ClOff}^n_\beta). \tag{4.2}$$

(I.e., $\lfloor x^s(y^i_1, ..., y^j_m) \rfloor^{n+1}$ is either in Abs^n_α or in Ral^n_α.) But then, it is also the case that $\mathsf{Tr}_\alpha(\lfloor x^s(y^i_1, ..., y^j_m \rfloor^{n+1})$ is in the extension of KKSat^n_α by clause 10 of definition 4.5. Moreover, $\lfloor x^s(y^i_1, ..., y^j_m) \rfloor^{n+1} \in \mathsf{Abs}^n_\alpha \cup \mathsf{Rel}^n_\alpha$: if only (4.1) or (4.1) and (4.2) holds, $\lfloor x^s(y^i_1, ..., y^j_m) \rfloor^{n+1} \in \mathsf{Abs}^n_\alpha$. If only (4.2) holds, $\lfloor x^s(y^i_1, ..., y^j_m) \rfloor^{n+1} \in \mathsf{ClOff}^n_\alpha$, and $\lfloor x^s(y^i_1, ..., y^j_m) \rfloor^{n+1} \in \mathsf{Rel}^n_\alpha$.

If, on the other hand, $\mathsf{Tr}_\alpha(\lfloor x^s(y^i_1, ..., y^j_m) \rfloor^{n+1})$ holds in c_α, i.e.,

$$\mathsf{KKSat}^n_\alpha(\lfloor \mathsf{Tr}_\alpha(x^s(y^i_1, ..., y^j_m)) \rfloor^{n+1}, X^{n+1}, Y^{n+1}, \varnothing)$$

or

$$\mathsf{KKSat}^n_\alpha(\lfloor \mathsf{Tr}_\alpha(x^s(y^i_1, ..., y^j_m)) \rfloor^{n+1}, \mathcal{M}^n_\alpha, M^n_\alpha, \bigcup_{\beta<\alpha} \mathsf{ClOff}^n_\beta).$$

then, by clause 10 of definition 4.5, $\lfloor x^s(y^i_1, ..., y^j_m) \rfloor^{n+1}$ is also in the extension of KKSat^n_α. Moreover, $\lfloor \mathsf{Tr}_\alpha(x^s(y^i_1, ..., y^j_m)) \rfloor^{n+1} \in \mathsf{Abs}^n_\alpha \cup \mathsf{Rel}^n_\alpha$ as above. Hence $\models^n_\alpha \mathsf{KF1}^n_\alpha$. The case for $\mathsf{KF2}^n_\alpha$ is analogous.

For the induction step, assume that $\mathsf{Tr}_\alpha(\lfloor \varphi \wedge \psi \rfloor^{n+1})$ holds in c_α, i.e.,

$$\mathsf{KKSat}^n_\alpha(\lfloor \mathsf{Tr}_\alpha(\varphi \wedge \psi)\rfloor^{n+1}, X^{n+1}, Y^{n+1}, \emptyset)$$

or

$$\mathsf{KKSat}^n_\alpha(\lfloor \mathsf{Tr}_\alpha(\varphi \wedge \psi)\rfloor^{n+1}, \mathcal{M}^n_\alpha, M^n_\alpha, \bigcup_{\beta<\alpha} \mathsf{ClOff}^n_\beta).$$

But then, KKSat^n_α applies to $\lfloor \varphi \wedge \psi \rfloor^{n+1}$, and hence KKSat^n_α also applies to both conjuncts, and so by induction hypothesis, KKSat^n_α applies to $\mathsf{Tr}_\alpha(\lfloor \varphi \rfloor^{n+1})$ and $\mathsf{Tr}_\alpha(\lfloor \psi \rfloor^{n+1})$, and hence KKSat^n_α also applies to $\mathsf{Tr}_\alpha(\lfloor \varphi \rfloor^{n+1}) \wedge \mathsf{Tr}_\alpha(\lfloor \psi \rfloor^{n+1})$. The reverse direction is analogous. Moreover, $\mathsf{Tr}_\alpha(\lfloor \varphi \wedge \psi \rfloor^{n+1})$ and $\mathsf{Tr}_\alpha(\lfloor \varphi \rfloor^{n+1}) \wedge \mathsf{Tr}_\alpha(\lfloor \psi \rfloor^{n+1})$ are both in $\mathsf{Abs}^n_\alpha \cup \mathsf{Rel}^n_\alpha$. Hence $\models^n_\alpha \mathsf{KF4}^n_\alpha$. The remaining cases are similar. □

Besides the fact that higher-order bicontextualism validates KF^n_α, it also validates a form of naïve truth introduction and elimination rules (see section 2.2.2). Truth elimination can be done inside a single context, i.e.,

$$\mathsf{Tr}_\alpha(\lfloor \varphi \rfloor^{n+1}) \models^n_\alpha \varphi.$$

which follows from clause 10 of definition 4.5. Truth introduction, on the other hand, might require a context shift:

Lemma 4.15.

$$\Gamma \models^n_\alpha \varphi, \text{ then } \Gamma \models^n_{\alpha+1} \mathsf{Tr}_{\alpha+1}(\lfloor \varphi \rfloor^{n+1}).$$

The key observation is that in context $c_{\alpha+1}$, we build the non-minimal fixed point that has all sentences that are valid in the previous context in its extension.

Proof. If $\Gamma \models^n_\alpha \varphi$, then $\lfloor \varphi \rfloor^{n+1} \in \mathsf{Abs}^n_\alpha \cup \mathsf{Rel}^n_\alpha$, and therefore, $\lfloor \varphi \rfloor^{n+1} \in \mathsf{ClOff}^n_\alpha$. Then, $\lfloor \varphi \rfloor^{n+1} \in \mathsf{ClOff}^n_{\alpha+1}$ because $\mathsf{ClOff}^n_\alpha \subseteq \bigcup_{\beta<\alpha} \mathsf{ClOff}^n_\beta$ which we use for the Z^{n+1}-variable in definition 4.8. Consequently, $\mathsf{KKSat}^n_{\alpha+1}$ applies to φ by clause 2 of definition 4.5, and then, by clause 10 of the same definition, also to $\mathsf{Tr}_{\alpha+1}(\lfloor \varphi \rfloor^{n+1})$. □

If, on the other hand, $\lfloor \varphi \rfloor^{n+1} \in \mathsf{Abs}^n_\alpha$, then we get truth introduction without a context shift:

Lemma 4.16. Let $\lfloor \varphi \rfloor^{n+1} \in \mathsf{Abs}^n_\alpha$. Then

$$\Gamma \models^n_\alpha \varphi \text{ implies } \Gamma \models^n_\alpha \mathsf{Tr}_\alpha(\lfloor \varphi \rfloor).$$

Proof. If $\lfloor \varphi \rfloor^{n+1} \in \mathsf{Abs}^n_\alpha$, then $\mathsf{KKSat}^n_\alpha(\lfloor \varphi \rfloor^{n+1}, X^{n+1}, Y^{n+1}, \emptyset)$ by definition 4.7, and hence, by clause 10 of definition 4.5, $\mathsf{KKSat}^n_\alpha(\lfloor \mathsf{Tr}_\alpha(\varphi)\rfloor^{n+1}, X^{n+1}, Y^{n+1}, \emptyset)$. □

This concludes the technical presentation of truth-theoretical bicontextualism for higher-order languages. In chapter 5, I'm going to develop bicontextualism for set theory before giving a philosophical discussion of both, truth-theoretical and set-theoretical bicontextualism in chapter 6.

4.4 OBJECTIONS AND REPLIES

Let me now consider two potential objections against the semantic theory developed above. The first one is from Rossi's original work on bicontextualism and goes back to Resnik [139]. Both argue that there is no higher-order quantification in natural language. The second is due to Picollo and Schindler, and goes back to Quine [130], who argue that second- and higher-order quantification can be mimicked by semantic predicates like truth.

4.4.1 *Absence of Higher-Order Quantification in Natural Language*

Rossi [144] argues that there is no higher-order quantification in natural language:

> [C]onsider $\forall x^1 \forall x^0(x^1(x^0))$. A literal reading of $\forall x^1 \forall x^0(x^1(x^0))$ would be nonsensical, as it would amount to something like "every x^1 x^1s every x^0", which is not an English sentence. That is, the x^1 would need to simultaneously be the syntactic object to which the quantifier applies *and* the predicate in the expression following the quantifier. Yet this doesn't seem possible in languages such as English. To be sure, plenty of English paraphrases of $\forall x^1 \forall x^0(x^1(x^0))$ are available, e.g. "everything has every property". However, this paraphrase is not genuine, as it effectively treats x^1 as a first-order variable, and not as a predicate. [144, p. 115, notation adjusted]

Following Rossi, I call higher-order quantification *genuine* if it cannot be paraphrased away by first-order quantification. Rossi's argument relies on a strong connection between the formal expression of a logical formalism and the surface structure of its natural language counterpart.[24] I agree that there is no natural language counterpart to $\forall x^1 \forall x^0(x^1(x^0))$ that respects the surface structure. But this does not mean that there is no higher-order quantification in natural language.

Rossi takes a predicate-based interpretation, i.e. an interpretation that uses the *grammatical predicate* to interpret the higher-order expressions, and then argues that there is no natural language counterpart corresponding to the predicate-based interpretation. However, the interpretation of higher-order expressions has been discussed extensively by many philosophers, and there are various possibilities.[25] For example, Boolos's plural interpretation of monadic second-order logic,[26] and Rayo and Yablo's interpretation of non-monadic second-order logic based on the non-nominal quantifier *somehow*.

---

24 In linguistics, the term "surface structure" refers to the actual, syntactic structure of a specific sentence, i.e., the form that one can see or hear. In contrast, the term "deep structure" is concerned with what a specific sentence expresses. For instance, the sentences "The dog chased the cat", and "The cat was chased by the dog" have different surface structures but the same deep structure. See Higginbotham [67].

25 See, e.g., Boolos [16, 12], Quine [130], Resnik [139], Shapiro [149], Rayo and Yablo [133], Williamson [169], Rayo [132], Florio [38], Rossberg [142], and Florio and Linnebo [42].

26 See section 3.1.3.

As described in more detail in section 3.1.3, according to the plural interpretation, we can interpret predicates such as $\text{red}(x^0)$ as "the red things" and second-order quantification such as $\exists r^1(r^1(a^0))$ as "there are some things such that a^0 is one of them". By the Rayo and Yablo approach, we interpret non-monadic second-order quantifiers such as $\exists r^1(r^1(a^0, b^0))$ as "somehow things relate such that a^0 is r^1-related to b^0."[27] Both approaches are controversial, however. Resnik [139] has argued that even in the face of plurals, phrases such as "the red things" seem to denote collection-like entities (i.e. sets). As for the non-nominal quantifier approach, Rossberg [142] has argued that problems arise in interpreting logical falsities such as $\exists r^1(r^1(a^0, b^0) \wedge \neg r^1(a^0, b^0))$. However, if there are different interpretations of the higher-order expressions available that are no predicate-based, we can't rule out higher-order quantification in natural language by ruling out only the predicate-based natural language reading of higher-order expressions.

Similarly, Higginbotham [68] has argued for higher-order quantification in natural language from purely syntactic considerations with his famous example "John is everything we wanted him to be",[28] which would not count as *genuine* higher-order quantification in the sense of Rossi.

In my view, all approaches to identify higher-order quantification in natural language by merely syntactic considerations must ultimately fail. After all, we can think of higher-order logic in two ways: either as Henkin-interpreted higher-order logic, or as full higher-order logic. Henkin higher-order logic is always reducible to multi-sorted first-order logic and is therefore not powerful enough to identify important mathematical structures up to isomorphism (due to the relevant metatheorems compactness and Löwenheim-Skolem).[29] Full higher-order logic lacks these metatheorems and therefore is powerful enough to identify important mathematical structures up to isomorphism and to make inferences that show the failure of compactness. However, these are semantic but not syntactic differences. The same applies to natural language: whether natural language quantification is true higher-order quantification (in the sense of full semantics), is a semantic, and not a syntactic question. Consequently, what the appropriate formalization of natural language quantification is cannot not be answered by syntactic considerations alone. Rather, we have to consider the semantics of natural languages as I have done in section 4.1.

4.4.2 *Mimicry of Higher-Order Quantification by Semantic Predicates*

Picollo and Schindler [122, 123] have argued that semantic predicates such as truth and satisfaction essentially mimic higher-order quantification in first-order languages. Following Quine, they identify "ascend to talk ... of sentences"[130, p. 11] as the key feature of the truth predicate in a first-order language. Consider, Leibniz's law formulated in second-order logic:

$$\forall v^1 \forall x^0 \forall y^0 (x^0 = y^0 \leftrightarrow (v^1(x^0) \leftrightarrow v^1(y^0))).$$

27 [133, p. 84, notation adjusted].
28 [68, p. 81].
29 See section 3.1.2 and [149, Sec. 6.2].

This can be mimicked in first-order logic, extended with a satisfaction predicate:[30]

$$\forall v^0(\mathsf{For}_{\mathcal{L}^1}(v^0) \rightarrow \forall x^0 \forall y^0(x^0 = y^0 \leftrightarrow (\mathsf{Sat}(v^0, x^0) \leftrightarrow \mathsf{Sat}(v^0, y^0)))).$$

Picollo and Schindler define a translation function τ from a second-order language to a first-order language with a truth predicate, and show that τ preserves inferences in the sense that all valid inferences from second-order logic are also valid inferences in first-order logic extended with an appropriate truth theory.[31]

The translation function τ can be generalized to higher-order languages, where the preservation of inferences also holds when an appropriate truth or satisfaction theory is added. Consequently, any instance of nth-order quantification for $n \in \mathbb{N}$ can be mimicked by a first-order language with appropriate semantic predicates.[32]

It is tempting to interpret their results as implying that higher-order quantification is unnecessary, since it can be mimicked by a first-order language with a truth theory. There seems to be a trade-off: ee can augment first-order languages with semantic predicates, or with higher-order quantifiers, but we don't need both. Linguistic parsimony combined with a Quinean skepticism about higher-order logic might lead one to adopt a truth theory instead of higher-order quantifiers.

However, abandoning higher-order quantifiers comes at the cost of expressive power. Suppose we are working in a countable first-order language $\mathcal{L}^1$ over a countable domain M. Then, the second-order domain $\mathcal{P}(M)$ is uncountable and contains elements that are not $\mathcal{L}^1$-definable. This causes a difference in meaning between a second-order formula and its first-order counterpart. For example, the first-order version of Leibniz' law expresses that two objects are identical iff they satisfy the same *definable* formulas. The second-order version, however, says that the two objects are the same iff they satisfy the same second-order variable where there is no definability restriction.

A skeptic of higher-order logic may not be convinced by this argument, as they may consider the restriction to definability to be unproblematic. But even for such a skeptic there is a strong difference between the first- and the second-order versions. The complete abandonment of higher-order quantification leaves us with an impoverished language, unable to speak categorically about *the* natural numbers. For whatever their formal counterpart of such an expression might be, it could only be formalized in a first-order language, and therefore could not unambiguously capture (up to isomorphism) important mathematical structures.

30 Picollo and Schindler [123, 881f.].

31 They use a theory of Tarski biconditionals for translations of higher-order sentences, which is a version of UTB (see [57, ch. 7]). Note that in section 4.4 of [123], the authors use Henkin semantics for second-order logic. This assumption is crucial for their Theorem 5 and Corollary 8, which essentially establish the preservation of inferences under the translation function τ.

32 The results can also be sustained by another line of reasoning: higher-order logic can be reduced to second-order logic where the higher-order predication relation is mimicked by a non-logical second-order relation constant ([149, sec. 6.2]). Using this technique, higher-order sentence can be reduced to their second-order counterparts, and then, using the translation function τ of Picollo and Schindler, to a first-order sentence with semantic predicates.

5

SET-THEORETIC BICONTEXTUALISM

I introduced bicontextualism as an intermediate position between absolutism and relativism in section 2.4. In chapter 4, I gave a formal development of the position in a truth-theoretical setting as a way to combine absolutism for non-expansion sentences with relativism for semantic paradoxes. As such, bicontextualism can be seen as an approach to the foundations of semantics, combining contextualism with absolute generality in a way that is compatible with the principle of semantic optimism (section 4.1.3).

In this chapter, I will develop the philosophical ideas of bicontextualism for the foundations of mathematics, namely for set theory. That is, I will develop a semantic theory, based on the RU approach, that can distinguish, at the sentence level, between inherently absolutistic sentences of the language of set theory, and inherently relativistic sentences of the language of set theory.

There are similarities and dissimilarities between set-theoretic and truth-theoretic bicontextualism. Both positions are similar insofar as they yield an intermediate stance between absolutism and relativism, namely the position that some sentences can—and should be—interpreted over an unrestricted 'domain', while other sentences cannot—on pain of contradiction—be interpreted over an unrestricted 'domain'. That being said, as in chapter 4, the main technical aim here is to define collections Abs and Rel which contain these sentences respectively.

The interpretation structure of truth-theoretic and set-theoretic bicontextualism differ significantly: in the truth theoretic case, each context shift is accompanied by a language expansion with additional expressive resources (the contextualised truth predicate). The new truth predicate had to be applied to all sentences of the previous contexts. So, each context c_α generated a new hierarchy of higher-order languages $\mathcal{L}_\alpha$ (definition 4.1) which required both, restricted *and* unrestricted interpretations. Because of that, I defined the absolutely unrestricted truths *relative to a context* c_α (definition 4.7).

In the set-theoretic case, the role of the contexts will be taken by the strongly inaccessible rank initial segments of the cumulative hierarchy, and the universal 'domain' will be the set-theoretic universe, i.e., the totality of all sets. While I will also work with language expansions, this will play a minor role. The only languages expansion is by the introduction of new terms. So, if $\mathcal{L}^1_\in(\alpha)$ and $\mathcal{L}^1_\in(\beta)$ are, respectively, languages of set theory with names for objects in V_α and V_β for $\alpha < \beta$, then $\mathcal{L}^1_\in(\beta)$ has just more atomic sentences of the form $a \in b$, where a and b are names for sets $a, b \in V_\beta \setminus V_\alpha$. Because of that, in the set-theoretic case, the new contexts—i.e., the new V_κ's—don't require a contextualisation of the inherently absolutistic truths. In other words, while in the truth-theoretic case, I ended up with a sequence of contextualised absolutely general truths, in the set-theoretic case, there will only be a single collection of absolutely general truths. But again, the main reason is that Parsons-Glanzberg

contextualism, according to Rossi's simplifications, operates by the introduction of new truth predicates.

The chapter is structured as follows: in section 5.1, I will give the motivation to develop a bicontextualist semantics for set theory. The main motivating reason will be that set theory is intended to be the theory about *all* sets, a task which theories are unable to achieve in standard model-theoretic frameworks. Nevertheless, this is not to say that the insights from the hierarchical cumulative-iterative conception of set should be neglected. Rather, the idea is that paradoxes highlight the limits of absolutism and the hierarchical nature of the underlying concept, while simultaneously revealing which sentences can be interpreted unrestrictedly. This, however, underlines the fundamental nature of the absolutely unrestricted truths as truth expressing core features of the set-concept. After the motivation, I'll give the presentation of set-theoretic bicontextualism, first, only heuristically (section 5.2), and then in full technical detail (section 5.3). The special status of absolutely unrestricted truths of the language of set theory will be taken up in section 5.3.2, in which I prove some basic facts about this collection of sentences. The philosophical discussion of set-theoretic bicontextualism, together with the philosophical discussion of truth-theoretic bicontextualism, will be given in chapter 6.

5.1 MOTIVATION

Intuitively, set theory should be the theory whose subject matter is the entire cumulative hierarchy, which comprises absolutely all sets. So, according to this view, the intended interpretation of the language of set theory has a 'domain' containing all sets.

As mentioned in section 2.1, sentences like 'Nothing is an element of the empty set', and 'Two sets are identical if and only if they have the same elements' express fundamental aspects of the set-concept. If only relativistic interpretations of these sentences are possible, this conflicts with their intended meaning. It is a fundamental fact about *all* sets—not only *some*—that they are not contained in the empty set, and that they satisfy the identity criterion given by the extensionality axiom.

However, as seen in chapter 2, such an interpretation cannot be given in standard model theory. I made this point before, but let me stress that the impossibility of an all-inclusive quantifier domain is especially critical for the case of set theory: according to model-theoretic semantics—which is arguably among the best formal account to semantics on offer—domains are sets. However, there cannot be a universal set on pain of contradiction, and hence, there cannot be a model-theoretic model compatible with absolute generality.

A natural way out would be a class-theoretic analogue of model theory that allows us to have classes as domains. On that account, the class of all sets V, which is the intended model of set theory, could serve as the domain of a class-sized model. In a reply to Parsons's "Sets and Classes" [116], George Boolos famously expressed a somewhat skeptic attitude:

> Wait a minute! I thought that set theory was supposed to be a theory about all, 'absolutely' all, the collections that there were and that 'set ' was synonymous with 'collection'. [14, p. 35]

The addition of a further layer of classes, Boolos continues, seems to force us into an iterative hierarchy of ever-growing collection-like entities accompanied by an iterative hierarchy of membership relations:

> [Class theory] cannot be the last word ... the considerations that prompt its adoption prompt the adoption of a still stronger theory. And of a yet stronger one. If one admits that there are proper classes at all, oughtn't one to take seriously the possibility of an iteratively generated hierarchy of collection-theoretic universes in which the sets which ZF recognizes play the role of the ground-floor objects? [14, p. 36]

A theory strong enough to model quantification over all sets would have to be developed with the aid of classes and a set-class membership relation. But then, what about collections of classes, and in particular, the collection of all classes? This collection cannot exist as a class on pain of (a class-theoretic equivalent of Russell's) paradox. This commits us to yet another layer, say *super*classes and a class-superclass membership relation. Again, we can consider collections of (all) superclasses which leads us to, say, *hyper*classes and a superclass-hyperclass membership relation, and so on.

These ideas have recently been taken up and spelled out in detail by Toby Meadows [103], who concludes that

> [this] approach to membership provides an intellectually honest account of where the standard classical tools of set theory lead us when we come to ask what kind of thing the collection of all sets is ... indefinite extensibility is not just witnessed by an ever-escaping ontology, it is also brought to life in the essential inability to provide a total theory of membership. [103, p. 1902]

The upshot, I believe, is that a natural setting in which the quantifiers of the language of set theory can range over absolutely all sets, is one developed with the RU approach and a non-objectual plural 'domain'. This approach allows us to give the set-theoretic quantifiers their intended meaning, and thereby also set theory its intended subject matter. At the same time, a plural interpreted metatheory avoids the commitment to an iterative hierarchy of larger and larger collection-like entities and membership relations. This blocks the pessimistic part of Meadows's conclusion concerning "the essential inability to provide a total theory of membership."

But, as also seen in chapter 2, not all sentences of the language of set theory can be interpreted unrestrictedly on pain of contradiction. In this sense, I think of the paradoxes as a limiting factor for which sentences of the language of set theory can, and which cannot, be interpreted unrestrictedly.

Nevertheless, it would be too hasty to give up on the project completely. With the philosophical method outlined in chapter 2, and spelled out in a truth-theoretical setting in chapter 4, it is possible to identify a collection of

sentences of the language of set theory that expresses truly absolutistic facts about sets.

The identification of such a collection is a task current set theory alone is unable to do: as we will see below, the universal model, having as its 'domain' the plurality of all sets, will be a model of ZFC. Moreover, the ZFC-axioms will be contained in the collection of *inherently absolutistic truths* of the language of set theory. However, ZFC alone cannot exhaust this collection, and so there remain inherently absolutistic truths of the language of set theory that are not provable in ZFC.

This collection is of a special status: it will contain all and only those sentences that are true and have no countermodels. As such, these sentences express some of the most fundamental features of the set-concept. No model of set theory, neither in the sense of an all-inclusive RU model, nor in the sense of structurally nice models of the cumulative hierarchy, falsifies the claims in this collection. Consequently, whatever property these sentences express, they are so fundamental that all sets have to satisfy them. As such, it is fundamental to investigate this collection in order to shed new lights on arguably the most important concept in the foundation of mathematics.

Besides the philosophical interest in the fundamental aspects of the set-concept that requires unrestricted quantification, there's a further philosophical issue with strict forms of relativism. As Sean Morris puts it

> [w]ith [the adoption of the cumulative-iterative conception of set and the abandonment of universal collections] we lost any mathematical investigation into what Cantor identified as the absolute infinite. [105, p. 4]

After the discovery of the Burali-Forti paradox in the late 19th century, Cantor changed his terminology from *the Absolutely Infinite* to *consistent* and *inconsistent multiplicities*.[1] As Hauser [63] argues, consistent multiplicities are just sets. Inconsistent multiplicities have been nicely summarized by Welch and Horsten:

> A multiplicity can be of such a nature, that the assumption of the "togetherness" ("Zusammenseins") of a multiplicity's elements leads to a contradiction, so that it is impossible to conceive the multiplicity as a unity, as a "finished thing." [Cantor] called such multiplicities absolutely infinite or inconsistent multiplicities. In latter day jargon we should call such "proper classes." [168, p. 94]

As described in section 2.2.1, the development towards the cumulative-iterative account in the decades following the discovery of the paradoxes established set theory as a mathematical theory of the infinite, including ordinal and cardinal arithmetic. However, this also lead to the neglect of the study of what Cantor initially termed the Absolute Infinite. But recently, philosophers have come to be more and more interested again.[2] Moreover, as Hauser [63] argues, Cantor's shift in terminology did not lead him to completely abandon the notion of the

1 See e.g., his famous letter to Dedekind from 1899, reprinted in [20, p. 443]. See also [168].

2 See Welch and Horsten [168], Boolos [14], Hauser [63], Rayo and Uzquiano [135], Weir [167], Morris [105], Uzquiano [160], and Linnebo [90].

Absolute Infinite. Rather, he thought it was impossible to assign order types and cardinal numbers to the totality of all sets. This does not mean that the totality of all sets do not exist together, but only that they do not exist together *as a set*.

At this point, the plural interpretation of such large collections of sets seems the natural way to go. However, another viable candidate is available:[3] Welch and Horsten [168] suggest a *mereological* interpretation of classes.[4] Nevertheless, I stay with the plural interpretation for two reasons: for one thing, Welch and Horsten rule out the plural interpretation because of the ontological innocence of pluralities. Since the main aim of their paper is to motivate a particular set-theoretic reflection principle, and since, on their view, reflection principles require *entities* that can be reflected, the plural interpretation clashes with their philosophical goals. For another, assuming mereological entities, even though it might be compatible with bicontextualism, seems to have no further advantage, but has the disadvantage of an unnecessary complex ontology as this would require to have sets, parts, and wholes, whereas on the plural interpretation, we only have sets and nothing else. So, what is unfavorable for the project of Welch and Horsten is very beneficial for the bicontextualist approach.

That being said, bicontextualism offers middle ground regarding semantics and ontology: on the one hand, bicontextualism draws the necessary conclusions from the paradoxes, i.e., that the concept set is essentially hierarchical, that unrestricted quantification is impossible for expansion sentences. But at the same time, bicontextualism allows us to have quantifiers ranging over the Absolute Infinite, taken to be the plurality of all sets.[5]

One might suspect a disanalogy between the truth-theoretic and the set-theoretic development of bicontextualism, assuming that the former rests on an inconsistent foundation. However, this is a misunderstanding. The contextualist truth theory employed here is not naïve; it avoids inconsistency through stratification via context-parameters. Therefore, the analogy between the two frameworks is both structurally and conceptually sound: both rely on hierarchical restriction mechanisms to sustain consistency while, in the bicontextualist framework, both allow a form of absolute generality.

5.2 HEURISTICS

The key idea of *bicontextualism for sets* is to decide sentence by sentence between an absolutistic and a relativistic interpretation: expansion sentences of the language of set theory, which are inherently relativistic, are interpreted relativistically in order to avoid paradoxes; by contrast, non-expansion sentences of

3 Of course, classes are also a possible candidate, but I have dismissed them already above.

4 See also Lewis [83].

5 Note that this approach also coincides with the Cantorian idea that the Absolute Infinite neither has an order type nor a cardinal number. The theory of order types is the theory of ordinals, and as such, it is embedded in set theory. Moreover, the theory of cardinals is a theory of growing infinities, and as such, it is also embedded in set theory. However, the plurality of all sets cannot be in the range of the set-theoretic quantifiers (because it does not form a set), and so it is not possible to assign it a cardinal or an order type.

the language of set theory, which are inherently absolutistic, are allowed to be interpreted over an all-inclusive 'domain' containing absolutely every set.

Crucially, I think, it is important not to presuppose the distinction between expansion and non-expansion sentences, but to construct a semantic theory for the language of set theory in such a way that the distinction actually drops out of the theory. In order to do so, I work with a plural 'domain' containing absolutely all sets for the absolute part. Since this 'domain' corresponds to the set-theoretic universe, I denote this universal plurality by V. For the relativistic part of the semantics, I work with a sequence of strongly inaccessible rank initial segment of the cumulative hierarchy, i.e., with the V_κ's. Since these ranks are sets, they are all contained in the universal domain.[6]

Formally, I use a language that has a name for every set—let $\mathcal{L}^1$ be such a language. Interpreting $\mathcal{L}^1$ over the universal 'domain' V results in an interpretation which is a total function. Interpreting $\mathcal{L}^1$ over a restricted domain, say some V_κ in V, results in an interpretation which is a partial function instead. The reason is that, since $\mathcal{L}^1$ has a name for every set, there are many more names in $\mathcal{L}^1$ than sets in V_κ (for any κ).[7] Names that do not correspond to sets in V_κ simply do not denote.[8]

I use a three-valued semantics with truth values $0, 1/2$, and 1. If the interpretation function is a partial function, the semantics will be such that every sentence with denotationless terms has value $1/2$, and sentences which only contain denoting names have classical truth values. The main thought is that according to the restricted models, sentences referring objects outside the current domain are nonsensical or off-topic.[9]

In the next step, I interpret all $\mathcal{L}^1$-sentences over and over again, over all the restricted domains (V_κ's) *and* over the all-inclusive, maximal plural 'domain'. This yields collections of sentences validated by all the different models, which allow me to define the inherently absolutistic and relativistic sentences of $\mathcal{L}^1$. More specifically, notice that an inherently absolutistic sentence cannot just be taken to be a sentence that is true according to the universal 'domain'. Think of the sentence $\exists\kappa\mathsf{IC}(\kappa)$, saying that there is an inaccessible cardinal. This sentence is false in all V-ranks below κ (assuming κ to be the first inaccessible). Hence, I will not count this sentence as inherently absolutistic, since it has at least one countermodel. I now take the inherently absolutistic sentences to be those that are true in the universal 'domain', and have no countermodel. The

6 Depending on the ambient set theory T, there are more or less V-ranks that are set-sized (i.e., T-provably exist). For instance, if $T = \mathsf{ZFC} - \mathsf{Inf}$ (the axiom of infinity), then only finite ranks exist. If, alternatively, $T = \mathsf{ZFC} + \exists!\kappa\ \mathsf{MC}(\kappa)$ (there exists exactly one measurable cardinal κ), than all V-ranks below κ exist. Either way, whatever provably exists according to the ambient theory T is contained as a set in the absolute 'domain'.

7 The use of large languages can be understood as a practical heuristic which does not require deep philosophical justification. If we extend the language of set theory $\mathcal{L}^1_\in$ to a larger language $(\mathcal{L}^1_\in)'$ by adding κ-many names, every $(\mathcal{L}^1_\in)'$-structure has a unique $\mathcal{L}^1_\in$-reduct. This means the semantic framework for such huge languages can always recover the $\mathcal{L}^1_\in$-fragment if required. However, the names serve a technical purpose, and so I include them in the semantics.

8 Note the slight abuse of notation here: since $\mathcal{L}^1$ has a name for every set, it cannot be a set itself. Hence, an interpretation of $\mathcal{L}^1$ can also not be a set. The reader should bear in mind that all these notions are pluralities.

9 See Beall [6].

inherently relativistic sentences are then defined via the inherently absolutistic sentences. To see this, consider, for some V_κ, the collection of all and only those sentences true in V_κ, and then subtract from it the sentences that are also true over the universal 'domain'. The resulting collection contains all and only those sentences true in V_κ *due to their relativistic nature*. The main claim, then, is that all the sentences of $\mathcal{L}^1$ that are classified as inherently absolutistic by the above theory are to be interpreted over the all-inclusive 'domain', whereas all the sentences that are classified as inherently relativistic are to be interpreted over restricted domains. The upshot is a semantic theory that allows me to reconstruct the elegant relativistic approach to the set-theoretic paradoxes within a framework that also contains an all-inclusive 'domain', over which, in spite of the paradoxes, absolutistic sentences can be interpreted.

5.3 TECHNICAL PRESENTATION OF SET-THEORETIC BICONTEXTUALISM

In this section, I give the formal details of set-theoretic bicontextualism. Unlike in the truth-theoretic case, the heuristics developed in section 5.2 require some reformulations of the definitions of the RU framework, mostly due to the fact that I'm working with models in the sense of partial interpretations (which leave out some names that don't denote sets in the current model). Moreover, I don't need the full power of the higher-order languages.

As mentioned before, I focus on models of set theory in the sense of the strongly inaccessible rank-initial segments of the cumulative hierarchy. The main reason to do so is that these models are largely accepted to be the intended interpretation of set theory because they perfectly encapsulate the cumulative-iterative concept of 'set'.[10] This significantly limits the scope of the object theory as it rules out first-order set theory due to the existence of many non-standard models (section 3.3.3). In order to avoid this, I work with the second-order theory ZFC_2 (definition 3.19).[11] As a consequence, in this section, I work with a second-order object language in a third-order metalanguage, more is not needed.[12]

In set theory, the language is usually stripped down to the bare minimum, containing just $\in$ as the only non-logical expression. As mentioned before, I allow adding names for every set which results in a very big language, to

10 See section 2.2.1.

11 Other options are possible. For instance, quasi-categoricity results can be obtained in a suitably adjusted version of ZFC in infinitary logic. A language $\mathcal{L}_{\kappa^+\kappa^+}$ which has conjunction and quantifier sequences of length κ^+ allows us, 1. to define an infinitary version of FOUNDATION that rules out infinite descending chains. (The usual first-order version only rules out first-order definable infinite descending chains, see Shapiro [149, p. 113 ff.] for details.) 2. it allows us to define formulae saying that the model contains κ-many elements. Then, it can be shown that any model satisfying that theory is isomorphic to $\mathcal{H}(\kappa)$, the class of all sets hereditarily of cardinality less than κ (Hanf and Scott [62] and Karp [77]).

12 Note that, as in section 4.1, the need to consider very large structures can give rise to the need to work with even stronger theories. As mentioned in footnote 14 of chapter 4, if a cardinal is not characterizable in a logic of order n, it becomes characterizable in a logic of order $n+1$. However, I leave this out for the moment and focus on the quasi-categorical axiomatizations of the cumulative hierarchy as given by ZFC_2.

be denoted by $\mathcal{L}^2_\in(\mathsf{V})$. I also allow to restrict this language to the strongly inaccessible rank-initial segments. Here is the definition of the object languages:

Definition 5.1. Let $\mathcal{L}^2_\in$ be a second-order language that satisfies the following conditions:

1. $\mathcal{L}^2_\in$ has $\in$ as its only relation symbol, and has no function symbols;
2. $\mathcal{L}^2_\in(\mathsf{V})$ results from $\mathcal{L}^2_\in$ by adding a name for each set;
3. for every every ordinal α, $\mathcal{L}^2_\in(\alpha)$ is the restriction of $\mathcal{L}^2_\in(\mathsf{V})$ to V_α.

I use the following notation: Con^0 is the set of first-order constants of $\mathcal{L}^2_\in$, Con^0_α is the set of first-order constants of $\mathcal{L}^2_\in(\alpha)$, and $\mathsf{Con}^0_\mathsf{V}$ is the set of first-order constants of $\mathcal{L}^2_\in(\mathsf{V})$. Since the variables of $\mathcal{L}^2_\in$, $\mathcal{L}^2_\in(\alpha)$, and $\mathcal{L}^2_\in(\mathsf{V})$, are the same I use Var^0 and Var^1 for these collections.

Having said that, I now turn to the definition of the RU model for the second-order language of set theory which has a name for every set. The model predicate has to be defined in such a way that it captures the case for $\mathcal{L}^2_\in(\mathsf{V})$ and for $\mathcal{L}^2_\in(\alpha)$, the restriction of $\mathcal{L}^2_\in(\mathsf{V})$ to V_α. This will correspond to the intuition given above that when we're using a name that denotes a set outside of the current model, this is just nonsensical or off-topic. Formally, this is implemented by letting the interpretation function, when used to interpret $\mathcal{L}^2_\in(\mathsf{V})$ over all sets, be a total function, and when restricted to smaller models, be a partial function. However, this makes it necessary to adapt the definition of the RU model. Usually, the RU approach is developed in such a way that it doesn't consider partial interpretation functions.[13] Formally, the restriction is carried out by requiring that only those constants in $\mathcal{L}^2_\in(\alpha)$ have denotation, and those in $\mathcal{L}^2_\in(\mathsf{V}) \setminus \mathcal{L}^2_\in(\alpha)$ have no denotation. For the limit case, where $\mathcal{L}^2_\in(\alpha) = \mathcal{L}^2_\in(\mathsf{V})$, $\mathcal{L}^2_\in(\mathsf{V}) \setminus \mathcal{L}^2_\in(\alpha) = \emptyset$, and hence there are no denotationless terms. Here is the definition:

Definition 5.2. For every unary type-2 variable X^2, the $\mathcal{L}^3$-formula "X^2 is an RU-Model of the language of set theory $\mathcal{L}^2_\in(\alpha)$", in symbols $\mathbb{M}_\in(X^2)$, is defined as:

$$
\begin{aligned}
\mathbb{M}_\in(X^2) :\leftrightarrow\ & \exists X^1(X^1 \sqsubseteq X^2) \wedge \Big[\exists x^0 X^1(\langle \ulcorner \forall \urcorner, x^0\rangle) \wedge \\
& \forall x^0\Big(X^1(x^0) \rightarrow \big(\exists y^0(x^0 = \langle \ulcorner \forall \urcorner, y^0\rangle) \vee \exists z^0 \exists y^0(x^0 = \langle z^0, y^0\rangle)\big)\Big) \wedge \\
& \forall x^0\Big(\mathsf{Con}^0_\alpha(x^0) \vee \mathsf{Var}^0(x^0) \rightarrow \exists! y^0\big(X^1(\langle x^0, y^0\rangle) \wedge X^1(\langle \ulcorner \forall \urcorner, y^0\rangle)\big)\Big) \wedge \\
& \forall x^0\Big(\big(\mathsf{Con}_\mathsf{V}(x^0) \wedge \neg \mathsf{Con}_\alpha(x^0)\big) \rightarrow \neg \exists y^0\big(X^1(\langle x^0, y^0\rangle)\big)\Big)\Big] \wedge \\
& \Big[\exists x^1 X^2(\langle \ulcorner \forall \urcorner, x^1\rangle) \wedge \forall x^0(X^2(x^0) \rightarrow X^1(x^0)) \wedge \\
& \forall x^1\Big(X^2(x^1) \rightarrow \big(\exists y^1(x^1 = \langle \ulcorner \forall \urcorner, y^1\rangle) \vee
\end{aligned}
$$

13 See Rayo and Uzquiano [135], Rayo and Williamson [136], Button and Walsh [18], Trueman [158], and Rossi [144] and section 3.4.

$$\exists y^1 \exists y_1^0, ..., \exists y_m^0 (x^1 = \langle y^1, \langle y_1^0, ..., y_m^0 \rangle \rangle)) \Big) \wedge$$
$$\forall x^0 \forall y^0 \Big(X^2(\langle \ulcorner \in \urcorner, \langle x^0, y^0 \rangle \rangle) \rightarrow X^2(\langle \ulcorner \forall \urcorner, x^0 \rangle) \wedge X^2(\langle \ulcorner \forall \urcorner, y^0 \rangle) \Big) \wedge$$
$$\forall z^1 \Big(\mathsf{Var}^1(z^1) \rightarrow \exists y_1^0, ..., \exists y_m^0 (X^2(\langle z^1, \langle y_1^0, ..., y_m^0 \rangle \rangle)) \wedge$$
$$\forall y_1^0, ..., y_m^0 (X^2(\langle z^1, \langle y_1^0, ..., y_m^0 \rangle \rangle) \rightarrow X^1(\langle \ulcorner \forall \urcorner, y_1^0 \rangle) \wedge ... \wedge X^1(\langle \ulcorner \forall \urcorner, y_m^0 \rangle))) \Big) \Big]$$

The definition is more or less identical to definition 3.34, properly restricted to a second-order object language. There is only one major difference, namely the following:

$$\forall x^0 \Big((\mathsf{Con}_{\mathsf{V}}(x^0) \wedge \neg \mathsf{Con}_{\alpha}(x^0)) \rightarrow \neg \exists y^0 (X^1(\langle x^0, y^0 \rangle)) \Big)$$

This is the part that takes care of interpretations that are partial: recall that when working in V_α, terms denoting sets outside V_α are non-denoting terms, and talking about those sets is nonsensical. The above clause does exactly that: if x^0 is a name contained in $\mathcal{L}^2_\in(\mathsf{V})$ but not in $\mathcal{L}^2_\in(\alpha)$, then the first-order part of the model contains no tuple that interprets x^0, and hence, x^0 is a denotationless term.

The definitions of the domain-component and of the variation of an assignment have to be suitably adjusted to a second-order object language:

Definition 5.3. Let $s \in \{0, 1\}$. For all type-$(s+1)$ variables Y^{s+1} and type-2 variables X^2, the $\mathcal{L}^3$-formula "Y^{s+1} is the type-s RU-domain of the $\mathcal{L}^2_\in(\alpha)$-RU-model X^2", in symbols $\mathbb{D}_\in(X^2, Y^{s+1})$, is defined as:

$$\mathbb{D}_\in(X^2, Y^{s+1}) :\leftrightarrow \mathbb{M}_\in(X^2) \wedge \forall x^s \Big(Y^{s+1}(x^s) \leftrightarrow \exists y^s (X^{s+1}(y^s) \wedge y^s = \langle \ulcorner \forall \urcorner, x^s \rangle) \Big)$$

Definition 5.4. Let $s, t \in \{0, 1\}$ and $s \neq t$. For any type-s variable x^s, and any type-2 variables X^2, Y^2,, the $\mathcal{L}^3$-formula "the $\mathcal{L}^2_\in(\alpha)$-RU-model Y^2 is a $\mathcal{L}^2_\in(\alpha)$-RU-$q^s$-variant of the $\mathcal{L}^2_\in(\alpha)$-RU-model X^2", in symbols $\mathbb{V}_\in(q^s, X^2, Y^2)$, is defined as:

$$\mathbb{V}_\in(q^s, X^2, Y^2) :\leftrightarrow \mathbb{M}_\in(X^2) \wedge \mathbb{M}_\in(Y^2) \wedge$$
$$\exists Z^1 (\mathbb{D}_\in(X^1, Z^1) \wedge \mathbb{D}_\in(Y^1, Z^1)) \wedge \exists Z^2 (\mathbb{D}_\in(X^2, Z^2) \wedge \mathbb{D}_\in(Y^2, Z^2)) \wedge$$
$$\forall x^t \forall z^t \Big(X^{t+1}(\langle x^t, z^t \rangle) \leftrightarrow Y^{t+1}(\langle x^t, z^t \rangle) \Big) \wedge$$
$$\forall x^s \Big[x^s \neq q^s \rightarrow \forall z^s \Big(X^{s+1}(\langle x^s, z^s \rangle) \leftrightarrow Y^{s+1}(\langle x^s, z^s \rangle) \Big) \Big]$$

Finally, in the definition of weak Kleene satisfaction, I take care of $\in$ directly:

Definition 5.5. The $\mathcal{L}^3$-formula "the $\mathcal{L}^2_\in(\alpha)$-RU-model X^2 with RU-domain Y^2 weakly Kleene satisfies φ", in symbols $\mathsf{WKSat}(\varphi, X^2, Y^2)$, is defined as: $\mathsf{WKSat}^1(\varphi, X^2, Y^2)$:iff

1. $\mathbb{M}_\in(X^2)$ and $\mathbb{D}_\in(X^2, Y^2)$ and $\varphi \in \mathsf{Sent}^2_\alpha$ and
2. $\varphi \equiv x^0 \in y^0$ and $X^2(\langle \in, \langle x^0, y^0 \rangle \rangle)$

3. $\varphi \equiv x^0 \notin y^0$ and $\neg X^2(\langle \in, \langle x^0, y^0 \rangle \rangle)$
4. $\varphi \equiv z^1(t_1^0, ..., t_m^0)$ and $X^2(\langle z^1, \langle t_1^0, ..., t_m^0 \rangle \rangle)$, or
5. $\varphi \equiv \neg z^1(t_1^0, ..., t_m^0)$ and $\neg X^2(\langle z^1, \langle t_1^0, ..., t_m^0 \rangle \rangle)$, or
6. $\varphi \equiv \neg\neg\psi$ and $\mathsf{WKSat}(\psi, X^2, Y^2)$, or
7. $\varphi \equiv \psi \vee \chi$ and $\mathsf{WKSat}(\psi, X^2, Y^2)$ or $\mathsf{WKSat}(\chi, X^2, Y^2)$, or $\mathsf{WKSat}(\neg\psi, X^2, Y^2)$ or $\mathsf{WKSat}(\chi, X^2, Y^2)$, or $\mathsf{WKSat}(\psi, X^2, Y^2)$ or $\mathsf{WKSat}(\neg\chi, X^2, Y^2)$, or
8. $\varphi \equiv \neg(\psi \vee \chi)$ and $\mathsf{WKSat}(\neg\psi, X^2, Y^2)$ and $\mathsf{WKSat}(\neg\chi, X^2, Y^2)$, or
9. $\varphi \equiv \forall z^s \psi(z^s)$ for $\psi \in \mathsf{For}^2_\alpha$, and for every W^2 s.t. $\mathbb{V}(z^s, W^2, X^2)$, $\mathsf{WKSat}(\psi(z^s), W^2, Y^2)$, or
10. $\varphi \equiv \neg\forall z^s \psi(z^s)$ for $\psi \in \mathsf{For}^2_\alpha$, and for some W s.t. $\mathbb{V}(z^s, W^2, X^2)$, $\mathsf{WKSat}(\neg\psi(z^s), W^2, Y^2)$.

As in the more general case (definition 3.43), the main difference between the strong and the weak Kleene satisfaction predicate is taken care of by the disjunction clause (clause 7.). A sentence $\psi \vee \chi$ is true if and only if (at least) one disjunct is true *and* the other disjunct has a classical truth value. Moreover, by the $\in$-clause (clause 2.), the above definition guarantees that only atomic clauses in which every term denotes get a classical truth value. Note that in the case where we interpret the language which has a name for every set, $\mathcal{L}^2_\in(\mathsf{V})$ over V, every name denotes, and so, by theorem 3.41, satisfaction is fully classical (the same holds when we interpret $\mathcal{L}^2_\in(\alpha)$ over V_α).

For the rest of this chapter, I adopt the following notational convention:

Remark 5.6 (Convention on the enumeration of inaccessible ranks)**.** The enumeration of V-ranks is from now on restricted to inaccessible ranks, i.e., by V_1 I mean *the first inaccessible rank*, by V_2 I mean *the second inaccessible rank*, ..., by V_α I mean *the αth inaccessible rank* of the cumulative hierarchy, and so on.

With the convention in place, fix a standard model V_α, i.e., the αth inaccessible rank of the cumulative hierarchy. This generates the collection of sentences that are true in V_α:

Definition 5.7. Let V_α be the αth inaccessible rank of the cumulative hierarchy. The collection of $\mathcal{L}^2_\in(\mathsf{V})$-truths of V_α, in symbols WKSat_α, is the extension of WKSat over V_α, defined for the restricted language $\mathcal{L}^2_\in(\alpha)$:

$$\mathsf{WKSat}_\alpha := \{\varphi \in \mathcal{L}^2_\in(\alpha) \mid \mathsf{WKSat}(\varphi, X, V_\alpha)\}$$

Similarly, we can define the collection of sentences that are true in the universe of sets V:

Definition 5.8. Let V be the universe of sets. The collection of $\mathcal{L}^2_\in(\mathsf{V})$-truths of V, in symbols $\mathsf{WKSat}_\mathsf{V}$, is the extension of WKSat over V, defined for the full language $\mathcal{L}^2_\in(\mathsf{V})$:

$$\mathsf{WKSat}_\mathsf{V} := \{\varphi \in \mathcal{L}^2_\in(\mathsf{V}) \mid \mathsf{WKSat}(\varphi, X, \mathsf{V})\}$$

Note that the weak Kleene satisfaction predicate as given in definition 5.5 becomes a fully classical Tarskian satisfaction predicate for the full second-order language $\mathcal{L}^2_\in(\mathsf{V})$ which contains a name for every set, when interpreted over the universe of all sets V:

Corollary 5.9. Let $\mathcal{L}^2_\in(\mathsf{V})$ be the full second-order language of set theory which contains a name for every set, and let V be the universe of sets. Then, the Tarskian and the weak Kleene definition of satisfaction coincide, i.e., for all $\varphi \in \mathcal{L}^2_\in(\mathsf{V})$

$$\mathsf{TSat}(\varphi, X, \mathsf{V}) \text{ if and only if } \mathsf{WKSat}(\varphi, X, \mathsf{V}).$$

Proof. Since in V, every term in $\mathcal{L}^2_\in(\mathsf{V})$ denotes a set, each sentence φ is either true of false, i.e., for each $\varphi \in \mathcal{L}^2_\in(\mathsf{V})$, LEM holds. We can therefore apply theorem 3.41 to get the desired result. □

This completes the presentation of the basics of the bicontextualist semantics for the language of set theory. In the following section, I will use this to identify the absolutistic and relativistic truths.

5.3.1 *Absolute and Relative Truths*

Let me now define the notion of *absolutely general truths* and *relativistic truths* of the full second-order language of set theory $\mathcal{L}^2_\in(\mathsf{V})$. Let us use the following abbreviations: let $\mathsf{WKSat}_\mathsf{V}$ and WKSat_α be the extension of the satisfaction predicate for $\mathcal{L}^2_\in(\mathsf{V})$ over the universe of sets V and the extension of the satisfaction predicate of $\mathcal{L}^2_\in(\mathsf{V})$ over V_α respectively as defined above. Generally, let $\overline{\mathsf{WKSat}_\alpha}$ be the complement of WKSat_α.

I start with the definition of the inherently absolutistic sentences. I first give the definition and unpack it afterwards. The *absolutely general truths* of $\mathcal{L}^2_\in(\mathsf{V}$ is defined as follows:

Definition 5.10. The *absolutely general truths* of the universal second-order language of set theory $\mathcal{L}^2_\in(\mathsf{V})$, in symbols $\mathsf{Abs}_\in$, is defined as follows:

$$\mathsf{Abs}_\in := \mathsf{WKSat}_\mathsf{V} \setminus (\mathsf{WKSat}_\mathsf{V} \cap \bigcup_{\alpha\in\Omega} \overline{\mathsf{WKSat}_\alpha})$$

The absolutely general $\mathcal{L}^2_\in(\mathsf{V})$-truths are certainly evaluated by the universe of set V, as it can be seen in the first part of the definition. However, we cannot just take these $\mathcal{L}^2_\in(\mathsf{V})$-truths because the universe of sets, under certain large cardinal assumptions, validates sentences such as 'There are 36 inaccessible cardinals', which are not validated by all ranks smaller than the 37th inaccessible rank of V. I have to subtract all the $\mathcal{L}^2_\in(\mathsf{V})$-sentences on which the universe of sets V and all the smaller models disagree. This is done in the second part of the definition: $\bigcup_{\alpha\in\Omega} \overline{\mathsf{WKSat}_\alpha}$ is the union of all complements of collections of $\mathcal{L}^2_\in(\mathsf{V})$-sentences that are true in some standard model. In effect, I subtract all sentences that are true in V, but false in one of the V_α's.

Let me now define the inherently relativistic truths. In this case, I cannot simply identify the inherently relativistic sentences with the relativistic truths

of some standard model, since this would only generate a subcollection of the inherently relativistic sentences. Thus, I start with the relativistic truths of some V_α, i.e., the sentences that are true *in a standard model* by virtue of their inherently relativistic character, and give the generalization to the whole collection of inherently relativistic truths afterwards. The *relativistic* $\mathcal{L}^2_\in(\mathsf{V})$*-truths of* V_α is then defined as follows:

Definition 5.11. The *relativistic truths of* V_α, in symbols Rel_α, is defined as follows:

$$\mathsf{Rel}_\alpha := \mathsf{WKSat}_\alpha \setminus \mathsf{Abs}_\in$$

Rel_α contains sentences of $\mathcal{L}^2_\in(\mathsf{V})$ which are true in the standard model V_α (the αth inaccessible rank of the cumulative hierarchy), but which are not inherently absolutistic. Consider, for example, the sentence 'There are 36 inaccessible cardinals'. This sentence is true according to the universe of sets (under appropriate large cardinal assumptions), and hence it is in $\mathsf{WKSat}_\mathsf{V}$. However, it is not contained in any WKSat_n for $1 \leq n \leq 36$, and so it is in, say, $\overline{\mathsf{WKSat}_{15}}$. Consequently, by the definition of inherently absolutistic truths, it is *not* in $\mathsf{Abs}_\in$. But from V_{37} on, the sentence becomes true, and so it is in WKSat_α and also in Rel_α for $\alpha > 36$. So we have, for every V_α, the inherently relativistic truths of V_α. To get the inherently relativistic truths *simpliciter*, we take the union of all the relativizations to standard models. The *inherently relativistic truths* of $\mathcal{L}^2_\in(\mathsf{V})$, in symbols $\mathsf{Rel}_\in$, is the union of all relativizations of inherently relativistic truths to standard models of the form V_α:

Definition 5.12. The *inherently relativistic truths* of the universal second-order language of set theory $\mathcal{L}^2_\in(\mathsf{V})$, in symbols $\mathsf{Rel}_\in$, is the union of all relativizations of inherently relativistic truths to standard models of the form V_α:

$$\mathsf{Rel}_\in := \bigcup_{\alpha \in \Omega} \mathsf{Rel}_\alpha$$

For example, $\mathsf{Rel}_\in$ contains all large cardinal sentences of the form 'There are α inaccessible cardinals' as discussed above, since they all have countermodels and enter Rel_β for $\beta > \alpha$. On this approach, then, large cardinal sentences of this form are considered to be relativistic sentences. This leads to the following classification of inherently absolutistic and inherently relativistic truths of $\mathcal{L}^2_\in(\mathsf{V})$:

Remark 5.13 (Bicontextualist classification of inherently absolutistic and inherently relativistic truths)**.** The bicontextualist semantics for the language of set theory classifies sentences as inherently absolutistic and inherently relativistic as follows:

- the *inherently absolutistic truths* are just the absolutely general truths of $\mathcal{L}^2_\in(\mathsf{V})$, as defined by $\mathsf{Abs}_\in$;
- the *inherently relativistic truths* are the union of all relativizations to standard models V_α, as defined by $\mathsf{Rel}_\in$.

Let me now turn to the bicontextualist treatment of the set-theoretic paradoxes. The treatment of Russell's paradox is exactly as one would expect and as outlined in section 2.2.1. That is, for each (not necessarily inaccessible) rank of the cumulative hierarchy V_α, the Russell set of V_α, R_α, is not a set of V_α, but enters the hierarchy only at the subsequent layer $V_{\alpha+1}$. In other words, the treatment of Russell's paradox by the bicontextualist semantics is the standard treatment of the cumulative hierarchy, and so bicontextualism is in line with the hierarchical treatment of paradoxes due to the hierarchical nature of the set-concept. This leads to the following picture, already presented in section 2.2.1 that visualizes the relation between Russell sets and ranks of V:

$$\begin{array}{ccccccccccccc} \ldots & V_\alpha & \in & V_{\alpha+1} & \in & V_{\alpha+2} & \in & \ldots & \in & V_\gamma & \in & V_{\gamma+1} & \ldots \\ & \cup| & \not\ni & \cup| & \not\ni & \cup| & \not\ni & & & \cup| & \not\ni & \cup| & \not\ni \\ \ldots & R_\alpha & & R_{\alpha+1} & & R_{\alpha+2} & & \ldots & & R_\gamma & & R_{\gamma+1} & \ldots \end{array}$$

Apart from that, there is no such thing as a universal Russell set on pain of paradox:

Lemma 5.14. There is no universal Russell set, i.e., no set R such that

$$\forall x(x \in R \leftrightarrow x \notin x).$$

Proof. If there were an R such that $\forall x(x \in R \leftrightarrow x \notin x)$, then $R \in R \leftrightarrow R \notin R$. Contradiction. □

Nevertheless, every set has a Russell set, and this statement is, in fact, an absolutely general truth of the language of set theory $\mathcal{L}^2_\in$:

Lemma 5.15. It is an absolutely general truth that every set has a Russell set, i.e., for every set A, there is a Russell sentence

$$\rho_A := \exists x \forall y(y \in x \leftrightarrow y \in A \wedge y \notin y)$$

such that $\rho_A \in \mathsf{Abs}_\in$.

Proof. Let A be a set, α be the smallest ordinal such that $A \in V_\alpha$, and $\kappa > \alpha$ be inaccessible. Then, $V_\kappa \models \mathsf{ZFC}_2$, and so by the axiom of separation, there is a set $R_A = \{x \in R_A \mid x \notin x\}$. Since $R_A \subseteq A$, and since V_α is transitive, $R_A \in V_\alpha$, so $V_\kappa \models \rho_A$. Since only languages of the form $\mathcal{L}^2_\in(\kappa)$ have names for A, the truth value of ρ_A is $1/2$ in all V_λ for λ inaccessible $< \kappa$, and from V_κ on, the truth value of ρ_A is 1. Moreover, R_A is a set, and hence it is among the sets, i.e., $R_A \preceq \mathsf{V}$. Hence, $\rho_A \in \mathsf{WKSat}_\mathsf{V}$. This implies that $\rho_A \in \mathsf{Abs}_\in$. □

Despite the fact that there is no universal Russell set, all the Russell sets are, of course, sets. This gives rise to a universal Russell plurality, RR, which is a subplurality of the universe of sets, i.e., $RR \preceq \mathsf{V}$, and which contains all and only Russell sets. In other words, for each set A, there is a Russell set R_A which is contained in RR, i.e., $R_A \preceq RR$.[14]

14 Another prominent set-theoretic paradox is the Burali-Forti paradox which arises when we consider the collection of all ordinals Ω. Working in naïve set theory, i.e., with naïve compre-

5.3.2 *Extracting Core Properties*

The class $\mathsf{Abs}_{\in}$ contains absolutely general truths of the language of set theory. These sentences are characterized by two features. First, every sentence $\varphi \in \mathsf{Abs}_{\in}$ is true according to the RU model that has all sets in its 'domain'. Second, no sentence $\varphi \in \mathsf{Abs}_{\in}$ has a counter-model in the cumulative hierarchy. Examples are the axiom of extensionality, self-identity claims, and the like.

A sentence φ which has these two features expresses fundamental properties of the set concept. Not only is it true in the all-inclusive RU model—i.e., $\varphi \in \mathsf{WKSat}_{\mathsf{V}}$—but also any level in the cumulative hierarchy large enough to decide φ makes it true—i.e., $\varphi \notin \overline{\mathsf{WKSat}_{\alpha}}$ for each α.

The question which sentences go into $\mathsf{Abs}_{\in}$ can be answered, at least partly, via absoluteness results. More precisely, Σ_1 upwards absoluteness and Π_1 downwards absoluteness (theorem 3.25) generalize to the absolutist case: whenever a Σ_1-sentence is validated by some V_α, it remains true not only throughout the hierarchy of V_α's, but also in the limit case where the model has as its 'domain' the universe V. Similarly, whenever a Π_1-sentence is validated by the absolutist model V, it remains true in every rank of the cumulative hierarchy. Note, however, that in the case of Π_1 downwards absoluteness, we cannot use the language $\mathcal{L}^2_{\in}(\mathsf{V})$, which has a name for every set. The reason is that if φ is a Π_1-sentence of $\mathcal{L}^2_{\in}(\mathsf{V})$ of the form $\forall x \psi(x)$, where ψ contains a term that some smaller language $\mathcal{L}^2_{\in}(\alpha)$ does not contain, then φ has truth value $1/2$ in V_α even though it has truth value 1 according to V.

The situation is somewhat different for Δ_0-sentences, since their absoluteness depends crucially on the transitivity of the sets, and the universe of sets, taken to be a plurality, is not transitive because pluralities are flat. To see this, consider $V_3 = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\{\emptyset, \{\emptyset\}\}\}$ which is transitive. However, if we only look at $\{\{\emptyset\}\}$, it is not transitive because it has an element, $\{\emptyset\}$, which is not a subset of $\{\{\emptyset\}\}$, because $\emptyset \notin \{\{\emptyset\}\}$. But of course $\{\{\emptyset\}\}$ is a set, and so it is in the universe of sets. This means that we cannot argue downwards from V by Δ_0 absoluteness. What we can do, however, is to argue upwards, because whenever a Δ_0-sentence is true in some V_α, the set to which the quantifiers are restricted is also in V, and so the sentence remains true. The other direction does not hold. All these considerations lead to the following propositions:

Theorem 5.16. Let φ be a Δ_0-formula of $\mathcal{L}^2_{\in}$. $\varphi \in \mathsf{WKSat}_{\alpha}$ for some α if and only if $\varphi \in \mathsf{WKSat}_{\mathsf{V}}$.

---

hension (see section 2.2.1), this collection would be a set. If so, it could be well-ordered and would therefore be isomorphic to a unique ordinal, since ordinals form a unique representation system of well-orderings. However, this well-order would have to be larger than any well-order contained in Ω, and therefore constitute a well-order (and with it an ordinal) not contained in Ω, contradicting the assumption that Ω is the set of all ordinals. The treatment of the Burali-Forti paradox is analogous to the treatment of Russell's paradox: for each rank V_α, the ordinals of V_α, Ω_α do not form a set in V_α. However, the collection ordinals of V_α do form a set at the subsequent rank $V_{\alpha+1}$ which is unparadoxical. This is the standard treatment of the Burali-Forti paradox in the cumulative-iterative account of set, and it is treated in the same way by the bicontextualist semantics. Consequently, according to bicontextualism, there is no set of all ordinals. But since ordinals are sets, there exists a plurality of all ordinals, $\Omega\Omega$ which is a subplurality of the universe of sets, i.e., $\Omega\Omega \preceq \mathsf{V}$.

Proof. This is a generalization of theorem 3.25.1. If φ contains no quantifiers (i.e., is atomic, or of φ is of the form $\psi \wedge \chi$, $\psi \vee \chi$, $\neg\psi$, or $\psi \to \chi$ where both ψ and χ contain no quantifiers), then $\varphi \in \mathsf{WKSat}_\alpha$ if and only if $\varphi \in \mathsf{WKSat}_\mathsf{V}$.

Let φ be $(\exists x \in y)\psi(x)$, and $y \in V_\alpha$. Assume for induction that $\psi(x) \in \mathsf{WKSat}_\alpha$ if and only if $\psi(x) \in \mathsf{WKSat}_\mathsf{V}$.

$\Rightarrow$: If $\varphi \in \mathsf{WKSat}_\alpha$, then $(\exists x \in y)\psi(x) \in \mathsf{WKSat}_\alpha$ for some $x \in V_\alpha$ such that $x \in y$. But then, since $V_\alpha \preceq \mathsf{V}$, x and y are in V. Hence, by induction hypothesis, $(\exists x \in y)\psi(x) \in \mathsf{WKSat}_\mathsf{V}$.

$\Leftarrow$: Assume $(\exists x \in y)\psi(x) \in \mathsf{WKSat}_\mathsf{V}$ for some y in V_α. Then, for some $x \in y$, $\psi(x) \in \mathsf{WKSat}_\mathsf{V}$. Since V_α is transitive, $x \in V_\alpha$, and so by induction hypothesis, we get $\psi(x) \in \mathsf{WKSat}_\alpha$ for $x \in V_\alpha$, and hence $(\exists x \in y)\psi(x) \in \mathsf{WKSat}_\alpha$. □

Corollary 5.17. Let φ be a Δ_0-formula of $\mathcal{L}^2_\in(\mathsf{V})$. If $\varphi \in \mathsf{WKSat}_\alpha$ for some α, then $\varphi \in \mathsf{WKSat}_\mathsf{V}$.

Proof. Immediate from theorem 5.16. □

Theorem 5.18. Let φ be of the form $\exists x\psi(x)$, where $\psi(x)$ is a Δ_0-formula of $\mathcal{L}^2_\in(\mathsf{V})$ with all free variables displayed. Then $\varphi \in \mathsf{WKSat}_\alpha$ implies $\varphi \in \mathsf{WKSat}_\beta$ for all $\beta > \alpha$, and $\varphi \in \mathsf{WKSat}_\mathsf{V}$.

Proof. This is a generalization of theorem 3.25.2. Let φ be $\exists x\psi(x)$, where $\psi(x)$ is a Δ_0-formula with all free variables displayed and assume that $\varphi \in \mathsf{WKSat}^\alpha$, i.e., $\mathsf{WKSat}(\exists x\psi(x), X, V_\alpha)$. Then, there is some $y \in V_\alpha$ such that $\mathsf{WKSat}(\psi(y), X, V_\alpha)$. Then, since $\psi(y)$ is Δ_0, by theorem 5.16, $\mathsf{WKSat}(\psi(y), X, \mathsf{V})$, and since y is a set and therefore in V, $\mathsf{WKSat}(\exists x\psi(x), X, \mathsf{V})$, i.e., $\varphi \in \mathsf{WKSat}_\mathsf{V}$. □

Theorem 5.19. Let φ be of the form $\forall x\psi(x)$, where $\psi(x)$ is a Δ_0-formula of $\mathcal{L}^2_\in$ with all free variables displayed. Then $\varphi \in \mathsf{WKSat}_\alpha$ implies $\varphi \in \mathsf{WKSat}_\beta$ for all $\beta < \alpha$, and $\varphi \in \mathsf{WKSat}_\mathsf{V}$ implies $\varphi \in \mathsf{WKSat}_\alpha$ for all α.

Proof. This is a generalization of theorem 3.25.3. Let φ be $\forall x\psi(x)$, where $\psi(x)$ is a Δ_0-formula with all free variables displayed and assume that $\varphi \in \mathsf{WKSat}_\mathsf{V}$, i.e., $\mathsf{WKSat}(\forall x\psi(x), X, \mathsf{V})$. Then, for each $y \preceq \mathsf{V}$, $\mathsf{WKSat}(\psi(y), X, \mathsf{V})$. Fix any V_α. Since $V_\alpha \preceq \mathsf{V}$ and since $\psi(y)$ is Δ_0, by theorem 5.16, $\mathsf{WKSat}(\psi(y), X, V_\alpha)$ for all $y \in V_\alpha$. But then, $\mathsf{WKSat}(\forall x\psi(x), X, V_\alpha)$, i.e., $\varphi \in \mathsf{Sat}_\alpha$. □

The above results can be used to get a clearer picture of which sentences are in $\mathsf{Abs}_\in$: since we have Σ_1 upwards absoluteness, all Σ_1-sentences that are true in the first inaccessible segment will remain true throughout the whole bicontextualist framework. Consequently, all Σ_1-truths of V_1 are in $\mathsf{Abs}_\in$. Moreover, since all Π_1-sentences that are true in the universe remain true no matter how far down the hierarchy we go, they also have no counter-models. Consequently, all Π_1-truths of V are in $\mathsf{Abs}_\in$. Finally, all Δ_0-sentences that are true in any V_α remain true in all other V_β's *and* in V, so they're also contained in $\mathsf{Abs}_\in$. So, if

$$A \subseteq_{\Sigma_1} \mathsf{WKSat}_\alpha, \qquad B \subseteq_{\Pi_1} \mathsf{WKSat}_\mathsf{V}, \qquad C \subseteq_{\Delta_0} \bigcup_{\alpha\in\Omega} \mathsf{WKSat}_\alpha$$

are the sets of Σ_1-sentences in WKSat_α, of Π_1-sentences in $\mathsf{WKSat}_\mathsf{V}$, and of Δ_0-sentences from all the WKSat_α's, then

$$A \cup B \cup C \subseteq \mathsf{Abs}_\in.$$

Moreover, also the ZFC_2-axioms are absolutely general truths of the language of set theory. This is a corollary of the following theorem:

Theorem 5.20. The universal domain satisfies the ZFC_2 axioms, i.e., for all $\varphi \in \mathsf{ZFC}_2$, $\mathsf{WKSat}(\varphi, X, \mathsf{V})$.

The proof consists in verifying that all ZFC_2-axioms hold in V. I write $\mathsf{V} \models \varphi$ for $\mathsf{WKSat}(\varphi, X, \mathsf{V})$. Most cases can be adopted from the proof of theorem 3.28 (that $V_\kappa \models \mathsf{ZFC}_2$). However, it is crucial to know that V, the universe of sets, is closed under set-of operations as encoded by ZFC_2 *by assumption*. This follows from the definition of V (section 3.3).

There is another important aspect: many steps in the proof of theorem 3.28 start by showing that if a particular set or sets are contained in some V_κ, then other sets are contained in V_κ as well. For instance, if $x \in V_\kappa$, then so is the powerset of x, its unionset, etc. Then, the proof proceeds by using Δ_0-absoluteness to show that not only is the particular set contained in V_κ, but V_κ also *thinks* that it contains the particular set, and therefore satisfies the axiom in question. This last step is crucial. It relies on the background assumption that the particular set we're looking for is contained in the background universe which, by assumption, has all the sets, is a cumulative hierarchy, and therefore is always right about what sets there are and how they are structured. Then, since the definition of these sets is Δ_0, it follows that the truth of the particular axioms in the background universe carries over to the V_κ. Now, theorem 5.20 operates directly on the background universe. This just makes explicit what is usually taken to be implicit in the proof. However, this has the direct implication that the second step—using Δ_0-absoluteness to show that the V_κ satisfies the respective axiom—is not needed anymore. Therefore, most steps of the proof aim to show that if some set is given, then its powerset, unionset, etc. is also a set. This implies that they are contained in the background universe, and so the background universe, which is always right about what sets there are and how they are structured, satisfies the respective axioms. Here is the proof:

Proof. EXTENSIONALITY Let $x, y \preceq \mathsf{V}$ be distinct. Then, there is a set $z \in x \wedge z \notin y$ (or *vice versa*), and since z is a set, $z \preceq \mathsf{V}$. Then,

$$\mathsf{V} \models \forall x \forall y (x \neq y \rightarrow \neg \forall u (u \in x \leftrightarrow u \in y))$$

which is equivalent to EXTENSIONALITY.

SEPARATION AXIOM Let $F \preceq \mathsf{V}$ and x be a set. Then, $F \cap x$ ($:\Leftrightarrow \{y \mid y \in x \wedge y \preceq F\}$) is a (possibly proper) subset of x and therefore a set. Hence $F \cap x \preceq \mathsf{V}$.

PAIRING Let x, y be sets. Since V is closed under set-of operations, $a = \{x, y\}$ is a set and therefore in V.

UNION Let x be a set. Then, all $y \in x$ are sets as well, and $z = \{y \mid y \in x\} = \cup x$ is also a set and therefore contained in V.

POWERSET If x is a set, then each subset $y \subseteq x$ is a set, and so is their collection $\mathcal{P}(x)$.

INFINITY $V_\omega \preceq \mathsf{V}$.

FOUNDATION The set-theoretic universe is well-founded by assumption.

REPLACEMENT AXIOM Let $G \preceq \mathsf{V}$, x be a set, and G be functional on x. Then, if y is the image of x under G, $|y| \leq x$, and so y is a set as well.

CHOICE The universe of sets is well-ordered by assumption.

□

Corollary 5.21. The ZFC axioms are absolutely general truths of the language of set theory.

Proof. By theorem 3.28, each inaccessible rank of V is a ZFC-model. Moreover, by theorem 5.20, the ZFC-axioms are in $\mathsf{WKSat}_\mathsf{V}$. Consequently, by definition 5.10, the ZFC-axioms are in Abs. □

5.4 OBJECTIONS AND REPLIES

Let me now consider two potential objections against the semantics for set theory developed above. The first one is based on the observation that forms of absolutism can be recovered in ZFC via reflection principles or via inner models. The second is that set-theoretic bicontextualism is too revisionary and might therefore be anti-naturalistic.

5.4.1 *Reflection Principles and Inner Models*

A first objection to bicontextualism for sets is that absolutism in set theory can be achieved via reflection principles or via inner models. Therefore, an extra absolutist interpretation is redundant, since truths about all sets can be extracted from reflection or inner models.

The main idea of reflection principles is the following: the cumulative hierarchy is so complex that it cannot be characterized by any formula φ of the language of set theory. The reason is that any such formula that is true in the universe of all sets is already true in some initial segment V_α. And so any attempt to characterize the cumulative hierarchy as the collection of all things that satisfy φ fails, because there is always a V_α such that φ characterizes V_α.

For example, V is closed under replacement and powerset, but there is some V_κ, for κ an inaccessible cardinal, which is already closed under replacement and powerset, and hence, closure under replacement and powerset is not a unique characterization of V. Gödel therefore concludes that "[t]he universe of all sets is structurally indefinable".[15] The truth of any such description is already *reflected* by some rank.

15 Wang [163, p. 280].

Since reflection principles are usually presented as biconditionals, they work both ways:[16] not only are truths about all sets reflected by initial segments, but also facts about the initial segments hold in the whole hierarchy. In this way, we can extract facts about the cumulative hierarchy, i.e., facts where the quantifiers can be taken to be unrestricted, from facts about initial segments.

Similarly, one could argue that absolutism can be mimicked by an inner model. Starting from the language of set theory ZFC, and a standard model V_κ of ZFC, an *inner model* is a definable transitive class $\mathcal{N}$ in V_κ, such that $\in^{\mathcal{N}} = \in^{V_\kappa} \restriction dom(\mathcal{N})^2$, $\mathcal{N} \models \mathsf{ZFC}$, and where $\mathcal{N}$ contains all the ordinals of V_κ. This last fact is crucial: Since inner models contain all ordinals, it is—depending on which inner model exactly we consider—at least close to absolutism. But if this is the case, then inner models at least allow us to mimic absolutism.

However, reflection principles and inner models do not quite provide what the bicontextualist wants, i.e., a maximally general interpretation of the set-theoretic quantifiers in the case of unproblematic sentences. Moreover, neither reflection principles nor inner models show that paradoxes do not force us into relativism. So even if absolutism can be achieved by such techniques, this does not mean that bicontextualism can be achieved by reflection or inner models.

5.4.2 *Anti-Naturalism*

Another line of criticism might arise from a naturalistically minded philosopher, and might look roughly as follows: Mathematicians work with ZFC (or some extensions thereof with large cardinal or forcing axioms), but they do not use bicontextualist semantics, and so neither should the philosopher.

I believe that, in the end, such criticism is misguided. There is no doubt that ZFC is the widely accepted canonical axiomatization of set theory, and I neither intend nor have the ability to change this. However, it is important to clarify what my proposal is, and what it is not.

Bicontextualism, at its core, classifies sentences in the language of set theory as either inherently absolutistic or inherently relativistic. The debate between absolutists and relativists is primarily a philosophical one, and my main goal is to contribute to this discussion. My use of languages with names for every set and the RU approach should be seen as methodological tools in pursuit of this philosophical aim.

Since the standard first-order language of set theory, $\mathcal{L}^1_\in$—which includes only $\in$ as a non-logical constant—is a sublanguage of my second-order language, $\mathcal{L}^2_\in(\mathsf{V})$, which contains a name to every set, we can easily recover the inherently absolutistic $\mathcal{L}^1_\in$-sentences from the inherently absolutistic $\mathcal{L}^2_\in(\mathsf{V})$-sentences. In essence, I am using a distinct methodological tool to achieve a philosophical goal, rather than challenging the mathematical canon. There is nothing anti-naturalistic about this approach.

What I explicitly do not intend is to settle inherently mathematical and set-theoretical questions such as 'What sets are there?'. This question can be reformulated as 'What is the ultimate large cardinal axiom to adopt?'. Answering this question is a task for set theorists. If, however, set theorists

16 See e.g., Devlin [27, pp. 25–6].

settle on a particular axiomatization, say $\mathsf{ZFC} + \mathsf{V} = \mathsf{UltL}$, then my approach can still be applied to identify the inherently absolutistic and the inherently relativistic sentences.

6

PHILOSOPHICAL ASPECTS OF TRUTH- AND SET-THEORETIC BICONTEXTUALISM

In this section, I will gather the philosophical insights to be drawn from the previous chapters. I will first consider questions that concern the semantic framework in general. More precisely, I will discuss two potentially problematic issues that have been discussed recently, and that can be raised against an absolutist-friendly semantics that incorporates higher-order expressions in general (section 6.1). This will yield a first approach to an absolutist-friendly interpretation of higher-order expressions (section 6.1.3). After that, I will draw some specific truth- and set-theoretic conclusions from the framework sketched above (section 6.2). Finally, I will compare the bicontextualist framework with other accounts in the literature (section 6.3). I will focus here on accounts that are somehow between absolutism and relativism. There are three alternatives to consider: recent work by Øystein Linnebo and James Studd [90, 84, 152] which builds on schematic and modal accounts of quantification, Kit Fine's *procedural postulationism* [35, 36, 37], as well as some accounts based on reinterpretations of the underlying language due to Williamson [170], Shapiro [148], and Uzquiano [160].

6.1 A FIRST STEP TOWARDS AN ABSOLUTIST-FRIENDLY INTERPRETATION OF HIGHER-ORDER LOGIC

In this part, I consider two objections that deliver valuable insights for the prospect of a consistent interpretation of higher-order logic that is compatible with absolutism. These objections are known as the *just more theory objection* and the *type problem*. I first introduce the just more theory objection (section 6.1.1). Then I introduce the type problem and discuss two possible solutions (section 6.1.2). Finally, I draw some conclusions from the discussion of the two objections and use these insights to make a first step towards an absolutist-friendly interpretation of higher-order logic (section 6.1.3).

6.1.1 *The "Just More Theory" Objection*

In Part B of their *Philosophy and Model Theory*, Button and Walsh present an interesting challenge to the prospect of using higher-order resources to single out isomorphism types of natural language expressions (or their formal counterparts), which is based on the referential indeterminacy of mathematical expressions.

Consider a formal language $\mathcal{L}$ and a $\mathcal{L}$-structure $\mathcal{M}$. We can construct a structure $\mathcal{M}'$ isomorphic to $\mathcal{M}$, in which each $\mathcal{L}$-expression has a different denotation, by using a permutation of the domain of $\mathcal{M}$ as the domain of

$\mathcal{M}'$.[1] Since isomorphism entails elementary equivalence, $\mathcal{M}$ and $\mathcal{M}'$ validate the same sentences even though they differ radically in $\mathcal{L}$-denotation. Since each domain gives rise to many permutations, we can construct many different isomorphic models with pairwise distinct $\mathcal{L}$-denotations.

Any attempt to label $\mathcal{M}$ superior to $\mathcal{M}'$ with respect to $\mathcal{L}$-denotation must be developed in terms of preferability (or similar notions). However, so the objection goes, if T is a theory of preferability, then T itself is prone to misinterpretation. After all, T is *just more theory*, and the permutation argument can be applied to any T-model to construct isomorphic copies which systematically misinterpret the language of T.[2]

We can rephrase the argument in terms of semantic theories. Let ST be your favorite set theory, which is used to develop the semantics of second-order logic. By the semantic clauses of full second-order semantics, a $\mathcal{L}^2$-sentence $\forall x^1 \varphi(x^1)$ is true in a model $\mathcal{M}$ iff every a^1 in $\mathcal{P}(M)$ satisfies $\varphi(x^1)$. Since $\mathcal{P}(M)$ is an (abbreviation of a complex) ST-expression, its meaning is given by the interpretation of ST. However, like the theory of preferability T, ST is *just more theory* and therefore prone to misinterpretation.

The key problem is that in order to secure the reference of the second-order expressions, we must first secure the reference of metalanguage expressions which rely on mathematically richer resources, in this case, set theory. However, such resources are *just more theory* and therefore might be misinterpreted. The question is, what, if anything, fixes the interpretation of expressions like *powerset*? Because this notion is essential for understanding of the full semantics. I will call this the *just more theory objection* (JMT).

The main source of JMT is that the interpretation of an object language is developed in a metalanguage containing rich mathematical resources which are themselves prone to misinterpretation.[3] In standard model theory, this involves *semantic ascent* from a formal language to a set-theoretically enriched natural-language metalanguage, for which we rely entirely on a completely unproblematic understanding of the metalanguage and the set-theoretical vocabulary it contains.

6.1.2 *The Type Problem*

The type problem of higher-order logic arises for absolutists when quantified expressions imply a commitment to genuine higher-order entities. This is evident in hierarchies of languages where expressions quantify over different types such as the languages used throughout this work. For example, consider the languages $\mathcal{L}^s$ and $\mathcal{L}^{s+1}$ and statements $\varphi_s := \exists x^{s-1} \psi(x^{s-1})$ and $\varphi_{s+1} := \exists y^s \psi(y^s)$. If we interpret these types as genuinely distinct entities, the truth of

1 See [18, Ch. 2] for details of the construction.

2 The objection goes back to Putnam and is discussed at length in [18]. See in particular their sec. 2.3 for references to Putnam's and other works on indeterminacy.

3 This is obvious for second-order logic, where we have set theory as the metalanguage. It is perhaps slightly less obvious for first-order logic, where we usually work in a natural-language metalanguage enriched with set-theoretic vocabulary. The mathematical richness of the metalanguage can be seen in quantifier satisfaction clauses: if $\mathcal{M} = \langle M, I \rangle$ is a model, we say that $\forall x \varphi(x)$ holds relative to $\mathcal{M}$ iff $\varphi(a)$ holds *for all* $a \in M$.

φ_{s+1} commits us to the existence of type-s entities absent in φ_s. Consequently, each new level in the hierarchy seems to require an expanded domain to include type-specific entities. If every $(n+1)$-order domain must contain entities absent from the n-order domain, then the truth of quantified statements leads to an ever-expanding ontological commitment, raising doubts about the prospects of any language in the hierarchy achieving absolute generality across types.

To tackle the type problem, two strategies have emerged: ontological innocence and restricted range of quantifier significance. The ontological innocence approach claims that higher-order expressions need not refer to genuine higher-order entities. For instance, Williamson's direct method [169] incorporates higher-order expressions directly into natural language without treating them as distinct entities, allowing for homophonic truth conditions without higher-order commitments. Alternatively, Boolos's plural interpretation [16, 12] frames higher-order quantification in terms of pluralities instead of postulating new entities. Both of these methods preserve the structure of higher-order logic while avoiding commitments to genuinely new higher-order entities.

The restricted range of quantifier significance strategy, developed independently by Schindler [147] and Florio and Jones [40, 39], builds on work of Russell [145]. This approach accepts the ontological commitment of higher-order logic but constrains it to meaningful ranges defined by type restrictions. While quantifiers range over an unrestricted 'domain', only specific entities can witness the truth of a quantified statement. For example, with a monadic predicate $z^n(x^s)$, z^n can meaningfully apply only to entities up to type $n-1$, making quantification meaningful only if the type restrictions are satisfied. This strategy allows higher-order expressions to coexist within a single, all-encompassing 'domain', that each type's quantifiers access, while placing restrictions on what forms a meaningful expression.

6.1.3 *A First Step Towards an Absolutist-Friendly Interpretation of Higher-Order Logic*

This section synthesizes prior insights to propose a speculative interpretation of the formal framework of the preceding chapters compatible with absolutism. The approach suggests viewing the open-ended sequence of formal languages $\mathfrak{L}$ (definition 3.31) as a partial analogue of natural language, drawing on *truth internalism* to avoid JMT.

Truth internalism has been proposed by Button and Walsh (see [18, Part B]) to avoid richer *external* metatheories, and instead constructs semantics solely by logical terms within a framework as the one proposed in this work. In the formal structure outlined in section 3.4, each language $\mathcal{L}^n \in \mathfrak{L}$ has its semantics within the next language, $\mathcal{L}^{n+1}$. This approach ensures that semantic operations occur within a hierarchically system, devoid of external meta-theoretic dependence. Extending Button and Walsh's approach to restrict truth internalism to mathematical vocabulary, I propose this purely logical framework as a general tool for understanding certain aspects of natural language semantics (though it does not fully capture complex linguistic phenomena like tense, modality, and indexicality). By taking $\mathfrak{L}$ as a (partial) formalization of natural language, the

result is a self-contained framework that provides both object languages and their semantics. This is a first step to circumvent the JMT objection.

A critical component for the absolutist aspect of the above framework is the plural interpretation of the first-order 'domain', which refrains from committing to higher-order entities. Plural quantification operates within the logical framework without introducing additional ontological commitments. Consequently, I take the plural interpretation of the all-inclusive first-order 'domain' as an indispensable feature of an absolutist-friendly interpretation of a higher-order framework. However, providing consistent interpretations for higher-order expressions requires careful consideration. In order to address the type problem, two main strategies are available, using ontologically innocent interpretations, or restricting the quantifiers ranges to specific types. Yet, relying on restricted quantifier ranges as a solution to JMT has its limits. If we misinterpret higher-order semantics from the outset, restricting quantifier ranges might worsen these interpretative challenges, as additional theory would be required to articulate the specific range limitations.

Whether the plural interpretation can fully resolve JMT for the interpretation of higher-order expressions remains uncertain. There's a possibility of misinterpreting higher-level plural terms just as we might misinterpret other mathematical expressions. Using a Williamsonian approach, we might want to incorporate plural expressions directly into natural language. If successful, their interpretation can be given homophonically in ordinary language without introducing new theory. This is not only true for plural interpreted higher-order logic. Likewise for genuine (relational or predicational) higher-order expressions, the Williamsonian approach addresses JMT effectively: incorporating *higher-orderese* as part of natural language aligns higher-order logic semantics with that of first-order logic in the sense that semantic clauses for higher-order logic can be given homophonically just like first-order semantic clauses. Consequently, the Williamsonian strategy requires no additional theory for the interpretations of higher-order logic and higher-order interpretations are only as prone to indeterminacy as interpretations of first-order expressions are. On that approach, sentences like $\forall x^1(x^1(t^0))$ can be interpreted as 'for any way of x^1-ing, t^0 x^1s', and this, in turn, can be formalized within $\mathfrak{L}$ by a type-2 predicate.

So far, I acknowledge truth internalism, the plural interpretation of the first-order 'domain' and treating higher-order expressions in a non-committing, ontologically neutral way. Thus, both JMT and the type problem objections are effectively avoided within this framework. But this leads to a significant question: what is the highest type in $\mathfrak{L}$? Rayo and Linnebo's [93] principle of semantic optimism (discussed in section 4.1.3) suggests that $\mathfrak{L}$ should at least contain all finite-orders. Furthermore, they propose a union principle which implies that the system can ascend to limit types such as $\mathcal{L}^\omega$, the smallest language containing all finite-order quantifiers. Given this, one might conclude that $\mathcal{L}^\omega$ provides a suitable representation of the upper limits of quantification in natural language. However, interpreting $\mathcal{L}^\omega$ as the formal counterpart of natural language presents challenges. If we assume that we actually speak $\mathcal{L}^\omega$, then by the principle of semantic optimism, natural language semantics must

be developed in a language $\mathcal{L}^{\omega+n}$, which poses a revenge-like problem for JMT, since such a solution would require a fully spelled out theory of ordinals.

In response, I argue for an interpretation of $\mathfrak{L}$ as an open-ended sequence of finite order languages rather than positing a last 'highest' language. This open-ended sequence allows each level of $\mathfrak{L}$ to be interpreted in terms of an ontologically noncommitting framework. Thus, higher-order expressions are treated with ontological neutrality. By framing $\mathfrak{L}$ as an open-ended hierarchy, interpretations of higher-order expressions remain flexible and extendable into new levels of semantic theorizing as necessary, thereby addressing the JMT objection and the type problem.

This leads to the following picture: an absolutist-friendly interpretation of higher-order languages must operate with an unrestricted plural first-order 'domain' and interpret all higher-order expressions ontologically neutral by homophonic semantic clauses within a higher-order metalanguage. This sequence of higher-order metalanguages must be open-ended so that, for every possible expression that might be added to the current language, it is possible to have a metalanguage predicate to interpret this expression. This approach takes languages and interpretations to yield a dynamic range of metalanguages of increasing types.

6.2 General conclusions concerning truth- and set-theoretic bicontextualism

In this section, I want to gather some general insights from the bicontextualist position developed in the preceding chapters.

It is worth noting that, despite the different subject matters of the two frameworks – truth and sets – both rely on the same plurality-based conception of semantics. In each case, the metatheory is based on a non-objectual plural language, and RU-models are defined as pluralities of ordered pairs consisting of expressions and entities of the domain. What differs is the order of the pluralities involved: the truth-theoretic framework, which accommodates higher-order object languages, requires a higher-order plural metatheory, whereas the set-theoretic framework, which remains second-order in its object language, can be formulated within a third-order metatheory, relying at most on superplurals.

To state the obvious (which has, in fact, already been said), bicontextualism implies that the relativistic argument from paradox is not as forceful as it is often assumed to be. This position demonstrates that we can have an unrestricted quantifier 'domain' without falling prey to paradox.

What is new, however, is that even though paradoxes can be avoided while maintaining an unrestricted quantifier 'domain', we can still draw the relevant conclusions from them. In particular, key concepts in semantics (such as 'truth') and mathematics (such as 'set') are *essentially hierarchical*. In other words, they are indefinitely extensible: they contain mechanisms that generate ever more instances of objects falling under these concepts, they enable diagonal arguments that allow us to transcend any given context, and these diagonal arguments reveal the limits of unrestricted quantification.

While it is perfectly reasonable to interpret quantifiers unrestrictedly in statements like $\forall x\ x = x$, $\neg\exists x\ x \in \varnothing$, and $\mathsf{Tr}_\alpha(\lfloor \forall x\ x = x \rfloor)$, the same does not hold for expansion sentences such as the Liar paradox and instances of naïve comprehension that would lead to Russell's paradox.

Let us examine the metaphysical picture underlying bicontextualism. In the truth-theoretic case, I worked with a 'domain' that encompasses absolutely everything, whereas in the set-theoretic case, I restricted myself to a 'domain' containing *only* all sets. This, however, is not truly a restriction: I could have considered a 'domain' including both all urelements and all sets, but for simplicity, I ignored urelements.

The broader metaphysical picture is roughly as follows: there exists a set-theoretic universe, V, which may or may not include urelements. Crucially, this 'domain' is not itself a set, yet it contains every set-sized domain. That is, in the purely set-theoretic case, it includes every rank-initial segment of the cumulative hierarchy; alternatively, if urelements are considered, it encompasses all such ranks built over an initial collection of urelements.

The number of sets contained in this universe depends on the underlying assumptions. For instance, if we assume only the most basic set-theoretic principles (excluding Infinity), then all sets are finite, and the universe V would effectively resemble what is usually denoted by V_ω. Within the bicontextualist framework, unrestricted quantification over all sets remains possible—even though standard model theory does not permit this. However, such a universe would be quite small compared to the one we get from more usual set-theoretic axioms. If, for instance, we assume strong large cardinal axioms—such as the existence of supercompact or extendible cardinals—the set-theoretic universe becomes vastly larger. Nevertheless, the bicontextualist framework still allows for unrestricted quantification over all sets that exist relative to a given background theory, such as $\mathsf{ZFC} + \mathsf{Sc}(\kappa)$, which asserts the existence of a supercompact cardinal κ. This means that whatever sets exist—beginning from κ and iteratively applying set-theoretic operations permitted by the other axioms—the bicontextualist framework allows quantification over their totality.

To state some more specific insights: The set-theoretic version of bicontextualism enables us to define a collection of inherently absolutistic sentences in the language of set theory, denoted by $\mathsf{Abs}_\in$. These sentences express irrefutable truths about the concept 'set'. By definition, they have no countermodels—neither in the sense of a rank-initial segment of the cumulative hierarchy nor in the broader sense of the entire set-theoretic universe. As such, they capture core structural features of the set-concept, which future research may further explore to gain deeper insights into the fundamental objects of mathematical ontology.

In contrast, the truth-theoretic version of bicontextualism is less concerned with truths about a specific domain and more focused on the behavior of the truth predicate itself. As such, truth-theoretic bicontextualism can be used to reveal fundamental features of the foundations of semantics, particularly within the hierarchy of languages outlined in preceding chapters. Following the principle of semantic optimism (section 4.1.3), this framework presents a highly dynamic and open-ended view of semantics. For example, in such a setting, we

can always introduce new expressions—not only as names for new objects but also as new semantic layers that express additional relational structures.

Moreover, within this framework, interpretations are treated on equal footing with relations. Essentially, as developed in section 3.4 and chapter 4, interpretations function as higher-order relations, linking linguistic expressions to the objects they refer to.

Of course, this is only a first step, and much remains to be explored. Future research may further investigate the connections between higher-order frameworks that align with absolutism on one hand and the principle of semantic optimism on the other. An open question concerns the broader implications for the foundations of semantics. Similarly, a thorough analysis of the inherently absolutistic truths in the language of set theory remains open for future research.

6.2.1 *Ideological vs. Ontological Commitment*

A useful distinction, introduced by Quine, is that between *ontological and ideological commitment*.[4] Roughly, ontological commitments concern what a theory says there is. As he famously put it, the ontological commitments of a theory are determined by the values of its bound variables. By contrast, ideological commitments concern the primitive notions a theory takes as given: the conceptual tools it uses to describe the world, such as predicates, logical constants, or modes of combination. While ontological commitment answers the question "What exists?", ideological commitment addresses "What must we say to describe it?"

This distinction plays a central role in evaluating the foundational virtues of logical and semantic theories. A theory might have minimal ontological commitments while still being ideologically rich—that is, it posits few or no objects but relies on a substantial expressive apparatus. Much of the appeal of a framework like plural logic lies in its promise of ontological parsimony. However, this comes at a potential cost in ideological complexity: new forms of expression may be required to simulate familiar notions (like quantification over relations or functions) without reifying them.

In the present framework, the introduction of the following primitive expressions increases the ideological commitment:

- $\langle \ulcorner \forall \urcorner, y^1 \rangle$: expresses that y^0 is in the quantifier domain; that is, y^0 is something over which we may quantify.
- $\langle z^0, y^0 \rangle$: expresses that z^0 denotes y^0.

These pairing operations allow us to simulate assignments, interpretations, and relational quantification internally, especially in contexts where the metatheory itself lacks full polyadic resources. However, this strategy marks an ideological commitment, particularly because it assumes pairing as a primitive mechanism. This is perhaps the most substantial ideological cost of the overall bicontextualist framework (or more precisely, of the RU approach).

4 See Quine [129, 128], and Cowling [23] for a more recent discussion.

That said, this commitment is not strictly unavoidable. As Button and Walsh argue in their *Philosophy and Model Theory*,[5] it is possible—at least in principle—to define the model-theoretic satisfaction relation for an nth-order object language within an $(n+3)$rd-order metalanguage, without relying on primitive pairing operations. On this approach, relations between variables, assignments, and formulas are captured via fully higher-order quantification. However, this comes at the expense of significant technical complexity, or, as they say, defining "satisfaction for second-order languages within a fifth-order deductive system ... is quite painful enough."[6] The gain in ideological commitment is, in my view, outweighed by the increased technicality and lack of clarity. For this reason, I assume the primitive pairing operations to be part of the basic toolkit.

6.2.2 *Flatness vs. Hierarchy: Two Perspectives on the Totality of Sets*

An important philosophical distinction arises between two ways of conceiving the totality of sets in the framework of set-theoretic bicontextualism: as a flat, plural 'domain' and as a structured collection of cumulative ranks. Both perspectives are based on the same totality—that is, the class of all sets—but are different lenses through which this totality can be studied.

On the one hand, the universal 'domain' is *flat*: it is a plural reference to all sets, taken collectively and without internal structure. In this perspective, all sets are equally part of the domain, and quantification over this plural totality is unrestricted. This perspective enables us to interpret sentences of set theory in a way that respects generality absolutism: a sentence is *inherently absolutistic* if it holds in this universal plural 'domain' and has no countermodel in any set-sized rank.

On the other hand, the ranks of the cumulative hierarchy—the V_k's—provide a *hierarchical* structuring of the very same sets. According to the cumulative-iterative conception of set, each set enters the hierarchy at a particular stage and depends only on sets from earlier stages. As Gödel famously puts it, "a set is something obtainable from the integers (or some other well-defined objects) by iterated application of the operation "set of" ".[7] The structure of the V_k's reflects this iterative nature: each stage V_k builds on earlier stages via the powerset and union operations. In this way, the cumulative hierarchy captures the iterative aspect of set formation.

Despite the difference in structure, these two perspectives are not in metaphysical tension. There are not *two kinds of existence* for sets. Rather, they are two conceptually distinct lenses through which the same totality can be understood. Every set that occurs at some level of the cumulative hierarchy also lives in the universal plural 'domain'. The flat perspective emphasizes unrestricted quantification, while the hierarchical perspective emphasizes the iterative and well-founded nature of sets. Both perspectives are essential: the plural 'domain' gives us logical breadth, while the cumulative hierarchy gives us structural depth.

---

5 See section 12.A, esp. fn. 25, pp. 286–287.
6 Button and Walsh [18, p. 286].
7 Gödel [53, pp. 474–5].

It might be tempting to identify the inherently absolutistic truths simply with those sentences that are true in the universal 'domain.' But this identification would be too coarse. Some sentences are true in the universal, plural 'domain,' yet fail to hold in some rank-initial segments. As mentioned earlier, sentences such as 'There are measurable cardinals' may be true in a universe that includes such large cardinals, but it has countermodels at every rank V_k that lacks them. As such, this sentence cannot be regarded as expressing a core feature of the set-concept. Inherently absolutistic truths, by contrast, are those that remain true throughout the cumulative hierarchy—they hold not merely in the flat 'domain', but also in all smaller approximations to it. It is this stability across the hierarchy including the flat, universal 'domain', that distinguishes inherently absolutistic from inherently relativistic, i.e., contingent or model-dependent, truths.

The ranks of the cumulative hierarchy can thus be viewed as successive approximations to the universal 'domain' of all sets. Each rank V_k captures a set-sized slice of the universe, closed under the standard set-forming operations up to that stage. As we move upward through the hierarchy, these slices become more inclusive, containing more of the set-theoretic universe. However, it is a crucial feature of this picture that no V_k, and indeed no set-sized collection of ranks, can ever fully exhaust the plural totality of all sets. The universal 'domain' remains forever open-ended relative to the hierarchy—always containing more than any set-sized model can capture. This is why the hierarchical perspective is not merely a technical artifact but a deep insight into the nature of the set-concept: sets are not just members of a static totality, but part of an iterative process. The inherently absolutistic truths are precisely those that emerge not from the height of the hierarchy but from their invariance across all of its stages.

Set-theoretic bicontextualism exploits both of these perspectives. While inherently absolutistic truths are those that hold in the flat plural 'domain' without being falsified in any V_κ, the hierarchical structure of the V_k's allows us to track how and why certain sentences fail to hold at particular stages—a feature that is crucial for consistency in the face of paradox.

6.2.3 *Direct vs. Indirect Treatment and Global Interpretability*

A disanalogy between the truth-theoretic and the set-theoretic versions of bicontextualism concerns the manner in which the core concept—truth or set—is treated within the object language itself. In the truth-theoretic case, the concept of truth is introduced explicitly: the object language contains an ordinal-indexed truth predicate (Tr_α), and the semantic theory provides a formal interpretation for this predicate. In this sense, truth is modeled directly. This reflects the standard treatment in light of the semantic paradoxes, which centrally involve self-reference and the use of semantic vocabulary. Accordingly, much of the technical machinery developed in Chapter 4—including Kripke fixed points, closing-off operations, and context-sensitive truth predicates—is designed to manage the use of truth predicates within the object language.

By contrast, in the set-theoretic case, no set predicate is introduced in the object language. Instead, the standard membership relation $\in$ plays the central

role, and the semantic theory interprets the axioms of second-order ZFC_2 relative to both restricted and unrestricted domains. Thus, the notion of 'set' is handled more indirectly: one interprets a theory about sets without formally using a set-hood predicate '$set(x)$' within the theory. This indirect approach mirrors standard practice in set theory, where the concept of a set is presupposed and theories such as ZFC describe how membership behave (via the axioms) rather than what sets are. Introducing a predicate '$set(x)$' would require a reformulation of ZFC_2.

An further apparent disanalogy between truth-theoretic and set-theoretic bicontextualism concerns the possibility of a global interpretation of the central notions in question. In the set-theoretic case, although the framework avoids a universal set and introduces no object-language predicate '$set(x)$', the membership relation $\in$ can be interpreted globally. This is achieved through the use of plural reference and the universal 'domain'. As a result, unrestricted set-theoretic quantification is possible, and the relation $\in$ can receive a global interpretation.

At first sight, the truth-theoretic case seems more limited. The language contains no global truth predicate; instead, each context has its own predicate Tr_α, and even though the extensions are build partly with reference to the unrestricted 'domain', these predicates express only the seemingly restricted notions of *truth-in-α* (as opposed to an unrestricted notion of *truth simpliciter*). This might suggest that the truth predicate is more context-dependent than the membership relation, and that truth resists global interpretation more fundamentally than set.

But this impression is misguided—or at least, it depends a lot on how one sets up the framework. The version of contextualism used here follows Rossi's simplification of Glanzberg's original proposal.[8] Glanzberg, in his original work, uses a single truth predicate and develops a technical machinery to model domain expansion.[9] In the version I use, each context comes with its own truth predicate, and these are interpreted individually. This simplifies the presentation considerably and, from the point of view of expressing the hierarchical structure of the notion of truth, doesn't lose anything essential.

However, this modeling choice does not imply that the concept of truth is more contextually restricted than the concept of set. The bicontextualist framework is flexible enough to accommodate Glanzberg-style single-predicate approaches if desired. But for the purposes, I followed Rossi [144] and took the approach with multiple truth predicates.

Importantly, this approach still allows for the formulation of contextually unrestricted truths. That is, within a given context α, certain sentences—such as $0 = 0$—can be declared true in an absolutely unrestricted sense. Similarly, set-theoretic bicontextualism allows for statements like $\neg\exists x \in \emptyset$ to be interpreted in an inherently unrestricted way. Thus, the apparent asymmetry in global interpretability disappears on closer inspection.

8 See Glanzberg [47] and Rossi [144].

9 Recall that Glanzberg's original work is inherently relativistic. The process of domain expansion is implemented implicitly, in the bicontextualist framework, in the sequence of relativistic closing-offs as described in section 4.3.3.

6.2.4 *The Hierarchical Nature of Truth and Set*

A central claim of this thesis has been that the concepts of truth and set are *essentially hierarchical*. This claim is motivated by the structural similarity between the semantic and set-theoretic paradoxes, which arise when we attempt to apply these concepts in an unrestricted or self-applying fashion. In both cases, diagonalization arguments show that certain forms of generality collapse into contradiction unless we impose constraints—typically, by introducing some form of hierarchical structure. The bicontextualist approach models this structure explicitly by indexing truth predicates and set-theoretic domains. This preserves the ability to express generality, while still respecting the limits imposed by the paradoxes.

But it is equally important to recognize that not all uses of truth or membership require semantic or set-theoretic ascent. Many applications of these concepts are stable or non-paradoxical—that is, they occur within a context and do not force us to step outside the current context. For example, asserting that $0 = 0$ is true, or that $\neg\exists x \in \emptyset$, does not demand any shift to a higher level or context. These are, in a sense, contextually unrestricted truths: they hold independently of further semantic or set-theoretic expansion.

This shows that the hierarchical nature of these concepts is not total. Rather, it emerges from unrestricted generality or self-reference as shown by the paradoxes. The framework developed here offers middle ground between flat conceptions (where no hierarchy is recognized and paradox looms) and strictly stratified approaches (where every use of truth or set is locked into a rigidly tiered system and absolutism is abandoned). In this respect, bicontextualism is not just a technical fix but a philosophical mediation between competing tendencies.

This perspective is similar to the one mentioned in section 6.2.2 (Flatness vs. Hierarchy), but it applies more broadly than to the set-theoretic case alone. The truth-theoretic variant of bicontextualism exhibits the same flexibility: it allows for expressions of general truth, while still being hierarchical as demanded by the paradoxes. What they show is not that the concepts of truth and set always require ascent, but that ascent must be possible, so that when it is needed, the framework can accommodate it.

6.3 COMPARISON WITH OTHER ACCOUNTS

Within this section, I will locate bicontextualism within the broader landscape of the absolutism–relativism debate. There are, broadly speaking, three different approaches to which I will relate bicontextualism.

First, there are the closely related *schematic* and *modal* accounts, most prominently developed by Øystein Linnebo [88, 90, 84] and James Studd [153, 154, 152], which I will discuss in section 6.3.1 and section 6.3.2 respectively. Second, there is Kit Fine's *procedural postulationism* [35, 36, 37] which I will discuss in section 6.3.3. Third, there are intermediate positions developed by Timothy Williamson [170], Gabriel Uzquiano [160], and Stewart Shapiro [148]

which share a common conceptual foundation. I will discuss these views in section 6.3.4.[10]

6.3.1 *Schematic Generality*

One of the most interesting alternatives to classical quantificational accounts of generality is based on a schematic understanding of quantification, or *schematic generality*.[11] As Linnebo [84] points out, the idea goes back to Russell and his famous distinction between 'any' and 'all':

> We can speak of any property of x, but not of all properties, because new properties would be thereby generated. [145, p. 230]

The reason for this generation is essentially a property-theoretic analogue of Russell's paradox that arises when we consider the totality of all properties. Russell's idea is that quantifiers involving 'all' (or 'every') presuppose a domain in form of a completed totality. In other words, the universal quantifier ranges over a domain of discourse, and so this domain has to exist (typically as a set) before we can define the semantic clauses for the interpretation of universally quantified statements. So understood, according to Russell, 'every' not only presupposes the existence of the domain as a completed totality, but also the existence of all the members of the domain. Recall that the existence of a single entity containing, for instance, all sets was exactly the problem that I identified causing Russell's paradox (section 2.2.1).

According to Russell, 'any' is different in that respect: the sentence 'Any dog that is well-trained behaves well' is not tied to a quantifier domain as 'Every dog that is well-trained behaves well' would be. Rather, the quantifier 'any' works via a free variable which can take any possible value. As such, schematic accounts of generality can be treated like a template that arises from a formula in combination with a specific side condition. For instance, a formalization of 'Every dog that is well-trained behaves well' in first-order logic would look as follows:

$$\forall x(\text{dog}(x) \wedge \text{well-trained}(x) \rightarrow \text{behaves-well}(x)).$$

If, instead, we use the 'any'-reading, we would rather get a schematic version in the spirit of the set-theoretic separation schema such as

> **Schematic Generality.** Any appropriate substitution for x in the following schema yields a true sentence of the underlying language
>
> $$\text{dog}(x) \wedge \text{well-trained}(x) \rightarrow \text{behaves-well}(x).$$

Together with a side condition specifying what can replace x to generate a correct instance of the schema, the above version of schematic generality allows

10 A further candidate mentioned earlier (section 5.1) is due to Toby Meadows [103]. This account develops an open-ended sequence of stronger and stronger notions of classes (classes, superclasses, hyperclasses, and so on), where unrestricted quantification over each of these collection-like entities is possible from the point of the stronger notion, but unrestricted quantification *tout court* is beyond reach.

11 The most important examples involve Studd [153, 152] and Linnebo [88, 90, 84].

us to generate infinitely many substitution instances. Moreover, there is no *prima facie* reason that the schematic reading is tied to reified quantifier domains (or any quantifier domains at all) or to set-theoretic size restrictions as in standard model-theoretic semantics. As such, the schematic account can be seen as a direct response to the main problem of absolutism caused by Russell's paradox, namely to the assumption that the universal 'domain' is an entity.

A particularly influential development of this idea appears in Hilbert's 1926 lecture "*Über das Unendliche*" (On the Infinite, Hilbert [69]), where he sets out the philosophical and methodological underpinnings of his finitist program. Hilbert proposes an interpretation of universal quantification that avoids ontological commitment to actual infinities. Instead of reading a universally quantified statement like $\forall x \varphi(x)$ as expressing a proposition about a completed domain of all objects, he interprets it as a rule stating that $\varphi(n)$ can be inferred for any numeral n.

Hilbert's motivation for this approach was grounded in his conviction that finite, elementary (arithmetical) mathematics is the only fully reliable part of mathematics. In contrast, infinite mathematics, such as set theory or real analysis, requires additional justification. Hilbert believed that such justification could be provided through consistency proofs, which ought to be carried out within finitary mathematics, demonstrating that no contradiction (e.g., $0 = 1$) could be derived in infinitary mathematics. As is well known, Gödel's incompleteness theorems had a profound and fundamentally limiting effect on Hilbert's program. However, Hilbert's ideas remain influential: consistency proofs are often formalized in systems like Primitive Recursive Arithmetic (PRA), which, following influential work by Tait [155], aligns with the kind of finitist reasoning Hilbert endorsed. In contemporary proof theory, however, such consistency proofs are typically carried out in systems that are strictly stronger than the theories whose consistency they aim to establish.[12]

Hilbert's approach replaces semantic generality with syntactic generality: what is asserted is not a truth about an infinite totality, but the provability of each instance. Hilbert emphasizes that generality is encoded in the use of schematic variables, which stand for arbitrary numerals. A universal statement, in this view, has meaning only in terms of the capacity to instantiate it at each numeral and derive $\varphi(n)$ using formal rules. Hilbert's approach is thus an early attempt of what we now call schematic generality: generalization not by quantifying over a domain but by applying a schema to arbitrary instances.

The schematic account has only limited expressive power. The step above that lead from the quantifier version of 'Every dog that is well-trained behaves well' towards the schematic generality version essentially proceeded by abstracting

12 In modern proof theory, consistency proofs are typically carried out by showing that a base theory such as PRA, extended by a suitable amount of transfinite induction (TI), proves the consistency of the target theory T. That is, one shows:

$$\mathsf{PRA} + \mathsf{TI} \vdash \mathrm{Con}(\mathsf{T}),$$

where $\mathrm{Con}(\mathsf{T})$ is a Π_1-statement expressing that no natural number encodes a proof of $0 = 1$ in T. This reflects Hilbert's requirement that the justification for infinitary systems be carried out using only finitist (i.e., PRA-acceptable) methods. For details, see Rathjen [131].

from the involved universal quantifier. No similar move can be done to abstract from an existential quantifier.

This limitation arises because existential quantification inherently relies on the assertion that at least one instance satisfying a given condition exists, whereas schematic generality, as outlined above, only specifies that any appropriate substitution results in a true sentence. That is, while the schematic approach provides a template for universally generalizable statements, it does not provide a means of asserting existence in the same way that $\exists x$ does in standard accounts of quantification.

To illustrate this point, consider the existential statement:

$$\exists x(\text{dog}(x) \wedge \text{well-trained}(x) \wedge \text{behaves-well}(x)).$$

This asserts that there is at least one well-trained dog that behaves well. A naive schematic analog would be something like:

> **Schematic Existence?** At least one appropriate substitution for x in the following schema yields a true sentence of the underlying language:
>
> $$\text{dog}(x) \wedge \text{well-trained}(x) \wedge \text{behaves-well}(x)$$

However, this formulation does not function as a true existential claim. It merely states that some substitution could produce a true sentence, but it does not provide a criterion for schematic existence as in the universal case. Unlike universal schematic generality, which allows an open-ended range of valid substitutions, existential quantification only postulates the existence of at least one verifying instance.

This can be further illustrated by the following observation: in the case of schematic generality, the dual nature of universal and existential quantification breaks down. A way to understand the scope of a schematic statement is by treating it as being effectively determined by the universal closure of the entire expression. That is, when we assert a schematic generalization such as

$$\text{dog}(x) \wedge \text{well-trained}(x) \rightarrow \text{behaves-well}(x),$$

this should be understood as standing for all appropriate substitutions of x, making it akin to the universally closed statement (without presupposing the existence of a domain)

$$\forall x(\text{dog}(x) \wedge \text{well-trained}(x) \rightarrow \text{behaves-well}(x)).$$

However, this has crucial logical consequences: in the case of schematic generalization, truth functionality cannot be guaranteed. Specifically, if we attempt to negate a schematic statement, we do not obtain the negation of a universal generalization, but rather a universally generalized negation. That is, the naive negation of our schema would yield

$$\neg(\text{dog}(x) \wedge \text{well-trained}(x) \rightarrow \text{behaves-well}(x)),$$

which, under schematic generality, would translate to

$$\forall x \neg(\text{dog}(x) \wedge \text{well-trained}(x) \rightarrow \text{behaves-well}(x)).$$

This is not equivalent to the expected negation of the universally quantified statement

$$\neg\forall x(\text{dog}(x) \wedge \text{well-trained}(x) \rightarrow \text{behaves-well}(x)),$$

which would instead amount to

$$\exists x \neg(\text{dog}(x) \wedge \text{well-trained}(x) \wedge \neg\text{behaves-well}(x)).$$

The crucial distinction is that, while standard first-order logic allows for an interplay between universal and existential quantifiers via negation (by De Morgan duality), schematic generality does not support this. Instead, the form of absolute generality available on the schematic account is restricted to universal claims and prevents existential statements from being derived by negation.

Slightly more technical, this limitation means that schematic generality allows us to express absolutely general Π_1-sentences (see section 3.3) but does not allow us to express absolutely general Σ_1-sentences. That is, while a schematic generalization can express a statement of the form $\forall x\varphi(x)$ it cannot capture statements of the form $\exists x\varphi(x)$. In sum, schematic generality, while powerful in avoiding domain assumptions, is inherently restricted to purely universal claims.

However, there is no obvious reason why absolutely general claims should face syntactic restrictions as the ones above. Certainly, a universally quantified and absolutely general statement is much more general than an existentially quantified one. However, the scope of what the sentences say is the same: the universally quantified statement says that absolutely everything that exists is such-and-such. The existentially quantified statement says that among all the existing things, there is at least one witness that has the property of being such-an-such. The similarity in scope arises from the fact that both claims essentially presuppose the same 'domain'—be it objectual or non-objectual. Only against the background of such an all-inclusive 'domain' do these claims express their intended meaning.

This highlights a crucial difference between schematic generality and bicontextualism. As I have showed in section 5.3.1 by a generalization of the Π_1 downwards absoluteness and the Σ_1 upwards absoluteness theorem, bicontextualism allows for absolutely general Σ_1 statements. In this respect, bicontextualism fares significantly better than schematic generality. Bicontextualism is not affected by syntactic restrictions on the complexity of true statements.[13]

13 Whether bicontextualism also allows for more complex statements such as Π_n or Σ_n claims (or perhaps even Π_n^m or Σ_n^m statements) in the collection of absolutely general truths is an open question. However, I see no general reason to the contrary as in the case of schematic generality.

6.3.2 *Modal Generality*

A way to overcome the expressive restriction of schematic generality is by spelling out the schemata using modal operators. This has lead to a vast amount of literature and to many very interesting and ongoing discussions.[14] It is certainly beyond the scope of this thesis to discuss all the different modal approaches to the philosophy of mathematics and the philosophy of language. I therefore focus on the ideas about modal generality developed in Linnebo [90] and Studd [152, ch. 6].

Let us consider the set-theoretic case. In a nutshell, the idea is the following: start from a formal language such as the first-order language of set theory. Add to this language the *primitive* modal operators $\Box$ and $\Diamond$.[15] Now, following Linnebo [90], it is entirely natural to consider the cumulative hierarchy as inherently potential. Whenever we are at some stage V_α in the cumulative hierarchy, there are some sets available for collection in V_α that have not yet been collected into a set. From this point of view, the set collecting together exactly those sets exists *potentially*. In other words, the as of yet uncollected sets potentially form a set at some later stage $V_{\alpha+1}$:

> Whenever a completed totality of objects is available, it is possible to extend the hierarchy to include a set with precisely these objects as elements. [90, p. 206]

When transitioning from V_α to $V_{\alpha+1}$, this potential existence becomes actual. This yields a very nice interpretation of the indefinite extensibility and the open-endedness of the set concept from section 2.2.1. Moreover, using modal plural logic, this can be described as follows:

$$\Box\forall xx\Diamond y\Box\forall u(u \in y \leftrightarrow u \prec xx). \tag{6.1}$$

A natural reading of eq. (6.1) is 'Necessarily, for any things, they possibly form a set'. Linnebo take this to be the only principle of set existence required for the main ideas of the cumulative hierarchy and for a justification of the ZFC-axioms.[16]

Linnebo's analysis is based on an extensionality principle analogous to Extensionality (section 3.3): sets are individuated by their elements. Once all

14 Early versions of modalised accounts of mathematics can be found in Reinhardt [137], Parsons [117], and Hellman [65]. More recently, the ideas have been taken up, in particular by Linnebo [88, 90], Berry [10], and Studd [153, 152]. Apart from the discussions directly involved with the inherently potential nature of mathematical objects such as the cumulative hierarchy (or maybe even the natural numbers as in Linnebo and Shapiro [94]), a further lively discussion concerns the notion of the potential nature of domains and intuitionistic logic (see in particular Crosilla and Linnebo [25, 24]) and the connection between predicativity and potentialism (see in particular Linnebo and Shapiro [96, 95]).

15 In the case of primitive modal operators, $\Box$ and $\Diamond$ are no longer interdefinable via $\Box\varphi :\leftrightarrow \neg\Diamond\neg\varphi$ or $\Diamond\varphi :\leftrightarrow \neg\Box\neg\varphi$. Since it is essential to consider the modal operators to be primitives, both have to be added to the language.

16 Assume we work with urelements, then, by eq. (6.1) there is a set of urelements. If we have no urelements, we still get a set of all things there are. But this time, this is just the empty set. Either way, the construction mechanism of the cumulative hierarchy is started and unfolds on top of this starting point by iterated applications of eq. (6.1).

elements of a set are available for collection, they get collected. This corresponds to the transition from potential to actual existence described above. If we have only reached the point were all elements of a set are generated (but nothing further), the existence of the set in question is only potential. However, the generation is immediate so that in the subsequent stage, the potential set becomes actual and exists from there on. Note that this implies that set existence is a stable principle: once a set has entered the cumulative hierarchy (i.e., once it became actual), it remains there and it will not get destroyed at later stages. In Linnebo's terminology, this means that as we go on to later stages, all previous stages will be fully captured by the later stage. Formally, this means just that for all ordinals α, $V_\alpha \subseteq V_{\alpha+1}$.[17]

Now, what has this modal analysis to do with absolute generality? Recall from section 6.3.1 that a schematic account of generality grants us some limited form of absolutism. It is possible to generalize over absolutely everything there is, but only when we restrict ourselves to Π_1-sentences. The modal analysis generalizes the schematic account. It corresponds to the schematic account regarding the interpretation of the indefinite extensibility and the open-endedness of the set concept. In the case of Russell's paradox, this simply means that for every sets there are, the collection of all non-self membered sets among them potentially form a set, that becomes actual at the subsequent layer. Since this pattern of reasoning can be repeated at any stage, every stage is potentially open for expansion, and consequently, the concept set is indefinitely extensible.

The modal account allows us to go beyond what was possible in the schematic account. Following Studd [152, Ch. 6], modalized quantifiers such as $\Box\forall$ and $\Box\exists$ can be used to mimic absolute generality without being committed to the existence of a universal domain, and without being committed to an absolutist understanding of the non-modalized quantifiers $\forall$ and $\exists$. The key is to understand the primitive modal operators. In my view, one of the most promising accounts on offer is Studd *interpretational account of modality*, according to which

> interpretational modal operators generalize about how the interpretation could admissibly be (holding the circumstances fixed). The truth of $\Box\psi$ depends on whether the proposition that would be expressed by ψ under other admissible interpretations of the lexicon is true (in the actual world). [152, p. 147]

In this sense, a modal statement of the form $\Box\psi$ is true, in the sense of Studd, if and only if, ψ is true on every admissible interpretation that results from the relativistic process of domain expansion. In some sense, then, $\Box\psi$ can be understood as 'However the lexicon gets expanded and interpreted by succeeding interpretations, ψ'.

17 There is, of course, much more to be said about the modal analysis of set theory. For instance, Linnebo gives a detailed analysis of the modal principles at play and establishes an extension of S4.2 together with some further stability principles. Moreover, he proves a translation function from the language of set theory into the modal language of set theory, together with a mirroring theorem saying that classical deducibility and modalized deducibility are preserved under his translation function, closing a (potential) gap between Linnebo's modal analysis of set theory and the more standard practice of mathematics.

Unlike the schematic account, modal generality can be combined with existentially quantified statement. On Studd's interpretation, $\Box\forall x\varphi(x)$ and $\Box\exists\varphi(x)$ can be paraphrased as

$\Box\forall x\varphi(x)$: However the lexicon of our language is admissibly interpreted, it will always be the case that everything within the domain of that interpretation satisfies the condition $\varphi(x)$;

$\Box\exists\varphi(x)$: However the lexicon of our language is admissibly interpreted, there will always exist at least one thing within the domain of that interpretation that satisfies the condition $\varphi(x)$.

Modal generality is not limited to Π_1 generality. In this respect, the modal approach fares better than schematic generality and is on the same level with bicontextualism.

The main difference between the modal accounts and bicontextualism lies in what primitive notions they assume. This has far reaching consequences. As Studd points out,

> the relativist's modal operator is on a par with the superplural or higher-order quantifiers deployed by the absolutist [152, p. 146]

Modal relativists (and absolutists) have to assume a primitive understanding of the modal operators, much like, following Williamson [169], absolutists have to develop a primitive understanding of higher-order and plural resources in order to give homophonic semantic clauses for sentences involving such resources.

My personal view is that developing a primitive understanding of higher-order and plural resources seems much less problematic then developing a primitive understanding of necessity and possibility. I do not take this to be an argument, but rather a personal judgment about where conceptual pressure is more acute. In my view, the process of domain expansion or the incorporation of further admissible interpretations as outlined by Studd presents us with a much more difficult challenge. Compared to that, adding plural resources to our language seems to me more natural and straightforward.[18]

Another point is that on the modal account (and also on the schematic account), generality is not quantificational. As we saw above, in the schematic setting, there is no quantifier domain over which we can interpret the unrestricted Π_1-sentences of schematic generality. In the modal setting, possible worlds semantics can only be a heuristic tool but must not be taken at face value. The reason is that spelling out possible worlds semantics requires a set containing all possible worlds. This, however, is not available for the account of modal generality as this would be tantamount to absolutism (since the quantifying over all possible worlds entails quantifying over everything there is and possibly can be). This really puts pressure on the development of a relativist-friendly primitive account of the modal operators. Consequently, assuming that we have

18 Studd's interpretational account of modality is not the only account of modality on offer. In particular, Linnebo [91] develops *dynamic abstraction* as a form interpretational modality. Nevertheless, many authors have critically discussed the primitive understanding of modalities, see [10, 74, 45, 3, 91, 152].

a sufficient relativist-friendly interpretation of the primitive modal operators, the generality expressed is not quantificational generality but a certain form of necessity. If we compare the interpretation of unrestricted generalities, we get the following analysis. For an absolutist, such a generality would simply be $\forall x \varphi(x)$ interpreted over an unrestricted 'domain'. The absolutist understanding of this claim would be 'For absolutely everything x there is, x is φ'. According to the modal account, such a generality would have the form $\Box \forall x \varphi(x)$, which they would interpret as 'Neccessarily, for anything x that can possibly exist, x must be φ'.

This highlights an open question for future research: it should be possible to define a translation function that maps necessary truths from the modal theory into the class of inherently absolutistic sentences of bicontexualism. Specifically, such a function would translate necessary truths from models at early stages in the cumulative hierarchy into the class $\mathsf{Abs}_{\in}$. Defining such a function would enrich the understanding of the relation between *necessity* in the modal framework and inherently absolutistic sentences in my framework. In particular, it might be worth to consider the following reflection principle, proposed by Linnebo [90]:

$$\varphi^{\Diamond} \to \Diamond \varphi \qquad (\Diamond\text{-Refl})$$

As Linnebo points out, standard reflection principles have a top-down character, whereas ($\Diamond$-Refl) reflects bottom-up: the truth of $\varphi^{\Diamond}$ can be interpret as the truth of (the modalized counterpart of) φ in the potential hierarchy of sets, and $\Diamond \varphi$ just states that φ—in which no modal operators occur—is possible. Understood that way, the quantifiers of φ range over all sets. A similar approach has already been suggested by Reinhardt [137], who also formulates a reflection principle that works bottom-up in a modal setting. Combining Reinhardt's and Linnebo's insights with a translation function from the modalized language into the language of bicontextualism, this might help to extract further insights about the class of inherently absolutistic sentences. Moreover, it might help shed light on the relation between these absolutistic truths and the involved notion of necessity that proponents of modal generality presuppose.

6.3.3 *Fine's Procedural Postulationism*

A version of relativism that is based on the idea that every universe is up for expansion is developed by Kit Fine in a series of papers [35, 36, 37] and is called *procedural postulationism*. The name is already suggestive. Consider a property $\varphi(x)$. If the current domain does not contain an object corresponding to this property (for instance, a set which is the extension of $\varphi(x)$), Fine interprets this as a *procedural postulate*: 'Introduce an object corresponding to $\varphi(x)$!', formalized as $!x\varphi(x)$.

In the case of Russell's paradox, this leads to the following reasoning: let M be a collection and consider the collection $R_M \subseteq M$ that contains all and only those objects in M that are non-self-membered. This leads to the well-known conclusion that R_M cannot be in M (section 2.2.1). Now, if M is the domain of the current context, R_M not only lies outside M, but outside the current

domain of quantification. According to Fine's procedural postulationism, we should expand the domain and introduce R_M. Studd [152, ch. 4.4] calls this the "Russell Postulate":

> **Russell Postulate.** Introduce an item which comprises the non-self-membered items of the initial universe!
>
> $$!y\forall x(x \in y \leftrightarrow \neg x \in x)$$

On Fine's view, paradoxes are patterns of reasoning that highlight the limits of a current domain, much like in my analysis in chapter 2. The essence of procedural postulates lies in their imperative nature: the reasoning of Russell's paradox is perfectly fine, and it shows us that there is some object, that we can make sense of, but that transcends the current domain, so introduce it! If successful, the postulate expands the initial domain and leads us to a more inclusive quantifier domain.

In a truth-theoretic setting, the Liar paradox would lead us to the introduction of further resources—be they expressive, syntactic resources or propositions—and so the procedural postulate can be compared to the context-shift mechanism of the contextualist. In the set-theoretic setting, the introduction of a Russell set for some rank V_α corresponds to the open-endedness of the set concept, since this step can always be done at any rank. Consequently, the procedural postulate might be interpreted as (part of) the motivation of going to higher and higher ranks in V. This analysis corresponds to what I have called *expansion sentences* at the end of chapter 2.

Fine is perfectly fine with having unrestricted quantifiers ranging over the current domain. However, every domain is open for expansion by the reasoning sketched above.[19] The question how exactly to understand the procedural postulates, and how they are triggered, is similarly difficult (and similarly open) to the question how to understand the context shift, and what causes it, of the contextualists.[20]

Fine and the bicontextualist agree—with many others—that paradoxes should be taken as limiting factors of the current quantifier domains, and that they highlight the hierarchical nature of the underlying concepts ('set', 'truth'). However, unlike Fine, the bicontextualist acknowledges that this does not necessarily mean that an all-inclusive 'domain' is unavailable, provided the notion of 'domain' is understood in an ontologically neutral way. In this sense, the bicontextualist accepts quantification over, say, the entire set-theoretic universe V.

This, however, means that for the bicontextaulist, procedural postulations cannot be accepted unrestrictedly. For instance, there cannot be a procedural postulation forcing the introduction of a universal Russell set on pain of contradiction. Because of that, the bicontextualist insists that procedural postulations are restricted somehow (e.g., to previously given sets).

19 See footnote 3 in chapter 2 on the distinction between restrictionist relativism and expansionist relativism.

20 See the references in footnote 46 of chapter 2 for further discussion on context shift and the underlying mechanisms.

Understood that way, what the bicontextualist considers to be the universal 'domain' is closed under accordingly restricted procedural postulates. This follows from lemma 5.14 and lemma 5.14. On the bicontextualist picture, it is an absolutely general truth that every set has a Russell set. However, since the universal 'domain' is not a set, there is no universal Russell set. Nevertheless, all the Russell sets are contained in the universal 'domain', so the universal 'domain' is closed under all Russell Postulates.

6.3.4 *Accounts due to Shapiro, Uzquiano, and Williamson*

The accounts proposed by Williamson [170], Shapiro [148], and Uzquiano [160] share a common approach: an indefinite sequence of reinterpretations of key semantic or set-theoretic terms. For this reason, I will discuss them together. The central idea in Williamson [170] and Uzquiano [160] is that, at any given time, we assign meanings to fundamental terms such as 'true,' 'false,' 'set,' and 'class.' However, when confronted with paradoxes such as Russell's paradox or the Liar paradox, we reinterpret these expressions, effectively shifting to a new language.

The underlying philosophical motivation is similar to the interpretational modality of Studd [152], but also to the shift of meaning that contextualists like Glanzberg encounter when transitioning from one context to another:

> We start with one set of correlative meanings for "say", "true" and "false"; we use them to construct a sentence that says nothing in that sense of "say"; but reflection on that sentence causes normal speakers to give "say", "true", and "false" a new set of correlative meanings, much like the previous ones except that the sentence in question says something in the new sense of "say"; the process can be repeated indefinitely. Normal speakers are not aware of the change, just as they are not aware of many ordinary processes of gradual change. They feel themselves to be going on in the same way, but they are not. [170, Sec. 6]

Like contextualists, Williamson postulates an open-ended sequence of reinterpretations of semantic concepts such as 'true' and 'false'. When we construct the Liar sentence λ_α within a given context c_α, the sentence expresses nothing in the sense of $\mathsf{Tr}_{\alpha+1}$ but has meaning only in the sense of Tr_α. However, this triggers a reinterpretation of the semantic vocabulary, expanding our semantic resources and allowing us to introduce $\mathsf{Tr}_{\alpha+1}$. Constructing $\lambda_{\alpha+1}$ then reiterates this process. This illustrates the inherently indefinitely extensible nature of the concept of truth.

Something similar can be said about expressions like 'set':

> For given any reasonable assignment of meaning to the word "set" we can assign it a more inclusive meaning while feeling that we are going on in the same way, and make correlative changes to the words in an iterative account of sets, to preserve it too. [170, Sec. 7]

Just as semantic notions involve a sequence of reinterpretations of semantic terms, the set-concept allows for an indefinite sequence of reinterpretations of the term 'set'.

Uzquiano [160] develops Williamson's ideas in the set-theoretic context, proposing a sequence of reinterpretations of $\in$ and 'set'. The idea behind the Williamson-Uzquiano account is that we can maintain an unrestricted 'domain' containing all sets in the background. However, the key set-theoretic notions ($\in$ and 'set') are repeatedly reinterpreted, but this process never results in an interpretation where they encompass all sets. In Uzquiano's terms, the reinterpretation of $\in$ and set corresponds to the following sequence:

$$\langle \mathsf{V}, \mathsf{set}_1, \in_1 \rangle, \quad \langle \mathsf{V}, \mathsf{set}_2, \in_2 \rangle, \quad \langle \mathsf{V}, \mathsf{set}_3, \in_3 \rangle, \quad ..., \quad \langle \mathsf{V}, \mathsf{set}_\alpha, \in_\alpha \rangle, \quad ...$$

where $field(\in_\alpha) = field(\mathsf{set}_\alpha) = V_\alpha$, i.e., $\in_\alpha$ is the restriction of $\in$ to some (strongly inaccessible) rank-initial sequence of the cumulative hierarchy V_α, and set_α is the set-predicate that covers just V_α.[21]

In this account, there is a universal 'domain' in the sense of generality absolutism, which is arguably a core tenet of absolutism. Moreover, 'everything' in this context means everything in the sense of a language-independent meaning, i.e., in the sense of an unrestricted 'domain'. However, this is only a partial victory for the absolutist. Expressions like 'set' or 'member' are still subject to restrictions based on the language currently in use, and this language is always open to reinterpretation, leading to a more inclusive language with a broader interpretation of 'set' or 'member'.[22] As Stewart Shapiro [148, p. 477] notes, this creates an asymmetry between the meanings of quantifiers on the one hand, and expressions like 'set' and 'member' on the other. The former can achieve absolute generality, while the latter cannot.

Williamson argues that inconsistencies arise once we assume that there is a 'super-meaning':

> The inconsistency is not in any one meaning we assign the iterative account; it is in the attempt to combine all the different meanings that we could reasonably assign it into a single super-meaning. [170, Sec. 7]

In other words, on the Williamson-Uzquiano account, the sequence of growing reinterpretations sequence never reaches a limit where we get a model of the form $\langle \mathsf{V}, \in \rangle$ in which $\in$ has its *true* meaning, i.e., where we really consider the extension of $\in$ over the whole set-theoretic universe V.[23]

Following Stewart Shapiro [148], these ideas can be further articulated using the distinction between first- and second-order quantification. Following Williamson [170] (and Zermelo [173]), Shapiro endorses unrestricted first-order quantification because, on his account, the formulation of Williamson (and

21 Note the similarity between this and Florio and Jones [40]. The range of significance, in Florio and Jones's terminology, is the extension of the respective reinterpretations of the $\in$-relation and the set-predicate.

22 This can be seen as a form of expansionism regarding set and truth in the sense of James Studd [152, ch. 4], combined with absolutism for quantifiers.

23 I do not specify an extension for set in the limit model because this is identical to V, the domain of the model.

Uzquiano) provides a way to articulate Zermelo's original ideas. According to Shapiro, Zermelo's original view works best against a background 'domain' containing all sets, where key expressions such as 'set' and 'membership' are open to reinterpretation. However, second-order quantification, understood as quantification over classes, is treated differently. According to Zermelo:

> what in one model appears as "ultrafinite non- or superset", is already a fully valid "set" with cardinal number and ordinal type in the next higher model and, in turn, serves itself as the [cornerstone (original: bed-stone)] in the construction of the new domain.[24]

On this view, first-order quantification can be unrestricted, following the Williamson-Uzquiano account. However, second-order quantifiers range over classes, and since the classes in one model become sets in a later model, this prevents an interpretation of 'class' that captures all classes.

On the bicontextualist account developed in the preceding chapters, such a 'super-meaning' is indeed available. In my view, Williamson's conclusion that a 'super-meaning' would lead to inconsistency is an example of the overly hasty conclusion derived from the relativistic argument from paradox (section 2.4.1). While it is true that a 'super-meaning' of a suitably adjusted version of the Liar sentence or of a sentence postulating the existence of a Russell set would lead to inconsistency, no such issues arise when interpreting the quantifiers in sentences like "Everything is self-identical' is true' or 'Nothing is an element of the empty set' with a 'super-meaning'.

In other words, within the bicontextualist framework, there is an interpretation of the quantifiers that allows them to range over absolutely everything, but this applies only to non-expansion sentences. Expansion sentences, by contrast, can never have an unrestricted interpretation of their quantifiers.

As for Shapiro's concern with second-order quantification, this issue simply does not arise in the bicontextualist account. In this framework, classes are not treated as entities but rather as pluralities that carry no ontological commitment (see section 3.1.3 and section 6.1). Therefore, second-order expressions, on the bicontextualist view, plurally denote sets, corresponding to (possibly proper) subpluralities of the set-theoretic universe V.

6.3.5 *Locating Bicontextualism*

All that being said, let me end this chapter with locating bicontextualism on a map reaching from strict relativism to strong absolutism. To the best of my knowledge, the strongest forms of relativism are found in Glanzberg's and Lavine's writings:

> no utterance can accomplish absolutely unrestricted quantification, in the sense of being interpreted as having its quantifiers ranging over a maximal and non-extensible domain of 'absolutely everything'. [49]

24 Modified from the original to replace the uncommon term "bed-stone" with the more familiar "cornerstone" for clarity. See Zermelo [173, p. 429 f.].

> it is not merely impossible to communicate the intention to quantify over absolutely everything, it is impossible to form such an intention. [81]

Whereas the strongest anti-relativism tendencies are found in Williamson's and McGee's writings:

> [relativism] is not a doctrine that can be propounded coherently. To uphold it, you would have to maintain that, for every discussion, there are things that lie outside the domain of that discussion. [101]

> generality-relativism blocks adequate semantic reflection in a meta-language ... [relativism] endangers the possibility of a reflective understanding of our own thought and language. [169]

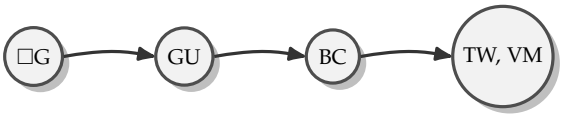
Whatever line we draw from relativism to absolutism, Glanzberg (MG) and Lavine (SL), as well as Williamson (TW) and McGee (VM) certainly correspond to the end points. This leads to the following picture, with the arrow pointing towards the more absolutist-friendly accounts, with boxes indicating relativistic positions and circles indicating positions that allow some form of absolutism:

Kit Fine (KF) allows for unrestricted quantification, but only restricted to the current context. So on that view, Fine is much closer to relativism than all the other accounts considered here as Fine does not allow for any from of true unrestricted quantification:

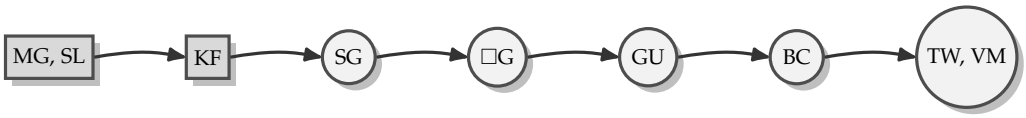
The schematic accounts of generality (SG), as developed (but not ultimately defended) in Linnebo [84] allows unrestricted quantification, but only for Π_1 sentences. The modal account of generality ($\Box$G) fares better in that respect as it does not have such syntactic limitations. Nevertheless, on the modal accounts, generality is not understood as quantificational generality, but as a form of necessity. Consequently, modal generality is not as absolutist-friendly as the positions of Williamson and McGee.

The accounts of (the early) Williamson [170], Shapiro [148], and Uzquiano [160] are not so easy to locate. Theoretically, they allow an unrestricted quantifier 'domain', but in particular the Uzquiano account does not lead to an unrestricted interpretation of important set-theoretic vocabulary that captures the intended interpretation of, say, the membership relation. I therefore locate this account as more absolutist friendly as the modal account, but not as absolutist-friendly as Williamson and McGee. I denote this account with GU for Gabriel Uzquiano because he presents the most well-developed account.

Bicontextualism (BC), in my view, is more absolutist-friendly than the other partial absolutist accounts KF (Fine [35, 36, 37]), SG and □G (schematic and modal generality), as well as the accounts due to (the early) Williamson [170], Shapiro [148], and Uzquiano [160] (GU). Bicontextualism allows for a non-objectual universal 'domain' as well as an quantifier ranging over this 'domain'. Moreover, it allows to evaluate sentences over this 'domain'. However, bicontextualism places some restrictions on which sentences to be evaluated over the unrestricted 'domain'. It is therefore less absolutistic than the accounts of Williamson and McGee. This leads to the following picture:



Put together, we get the following map of the absolutism-relativism debate, ranging from strict relativism in the sense of Glanzberg and Lavine, over Fine's procedural postulationism, to schematic and modal generality, to the picture developed by Uzquiano, to bicontextualism, and finally, to strong absolutism in the sense of Williamson and McGee:



7

CONCLUSION

In this work, I have developed an intermediate position between generality absolutism and generality relativism. I introduced both relativism and absolutism, outlined the debate between their proponents, and identified bicontextualism as a middle ground that combines the best aspects of both (chapter 2). In the subsequent chapters (4 and 5), I developed the bicontextualist position in both truth-theoretical and set-theoretical contexts. Additionally, I discussed the philosophical implications of this position, derived insights for a consistent interpretation of higher-order expressions compatible with generality absolutism, and situated bicontextualism within the current debate (chapter 6).

In the preceding chapters, I set myself several goals, and it is now time to assess whether I have achieved them. First, the intermediate position I developed should be:

1. a more fine-grained position in the absolutism-relativism debate.

Regarding the absolutist aspect of the semantics, it should be sufficiently absolutistic to:

2. allow for kind generalizations;
3. respect the intended interpretation of prima facie examples; and
4. avoid the relativistic disadvantages related mentioned in section 2.1.

Furthermore, bicontextualism should preserve the key insights derived from the relativistic solutions to paradoxes. Specifically, the bicontextualist's treatment of paradoxes should be analogous to the relativistic approach, meaning it should:

5. maintain the hierarchical structure of relativistic solutions to paradoxes (and with it, the hierarchical nature of the concepts 'truth' and 'set').

The bicontextualist position should also:

6. split the truths of the underlying language into *inherently absolutistic* and *inherently relativistic* truths; and
7. allow us to argue, contra Dummett, that paradoxes do not compel us to adopt a strict form of relativism.

Let me now briefly assess the extent to which these goals have been achieved: Bicontextualism indeed offers a more fine-grained position in the absolutism-relativism debate. Unlike strong forms of absolutism, it prevents an interpretation with unrestricted quantifier 'domain' for expansion sentences. Yet, unlike strong relativism, bicontextualism does not rule out an unrestricted quantifier 'domain' *tout court*, allowing non-expansion sentences to achieve absolute generality (see point 1). Since bicontextualism allows unrestricted quantification, kind

generalization can be achieved via quantifier restrictions to specific predicates (e.g., 'is a donkey'). However, unlike relativism, bicontextualism can articulate these quantifier domain restrictions with the help of the universal 'domain' in the background (see point 2).

Moreover, bicontextualism respects the universal character of prima facie absolutistic examples, which are typically classified as non-expansion sentences, and grants their quantifiers unrestricted range (see point 3). In the bicontextualist framework, semantic notions like satisfaction and logical consequence are defined universally, avoiding the issues that can come up in relativistic approaches. However, unlike relativists, these advantages do not come at the expense of semantic pariahs (see point 4).

The hierarchical structure of relativistic solutions to paradoxes has been successfully implemented in both truth- and set-theoretical contexts. Therefore, bicontextualism is not revisionary; it respects the insights of recent philosophical and mathematical investigations and validates the hierarchical nature of concepts like 'truth' and 'set'. Unlike relativism, however, bicontextualism is powerful enough to combine this with an unrestricted reading of quantifiers for some sentences (see point 5).

Bicontextualism crucially divides the sentences of the underlying language into *inherently absolutistic* and *inherently relativistic* truths. Importantly, bicontextualism does not presuppose this distinction but develops the semantic theory in such a way that the distinction naturally emerges. This is an important step towards understanding the fundamental truths in both semantics and mathematics (see point 6). Finally, bicontextualism synthesizes the relativistic treatment of paradoxes with generality absolutism. This implies that paradoxes alone do not justify adopting a strict form of relativism (see point 7).

Moreover, I emphasize that the truth-theoretic part of this thesis does not rely on naïve or inconsistent theories of truth. The contextualist theory employed is stratified and avoids paradox through regulated context shifts. As such, there is no foundational disanalogy between the truth-theoretic and the set-theoretic implementations of bicontextualism. Both maintain consistency through hierarchical structuring and rely on a unified way to resist the relativist's argument from paradox.

A central question posed in Chapter 1 was whether the paradoxes of truth and set force us to abandon generality absolutism. The answer provided by bicontextualism is now clear. The paradoxes indeed reveal limits to generality, but not of a global kind. What they show is that certain sentences—namely, expansion sentences—cannot be interpreted in an absolutely unrestricted way. These are precisely the kinds of sentences that involve paradox-generating constructions as in Russell's paradox or the Liar paradox. However, non-expansion sentences, which do not produce such effects, can be interpreted over a plural, unrestricted 'domain'. Bicontextualism therefore refines the relativist lesson: paradoxes are not an argument against absolutism as such, but a constraint on which sentences can receive an absolutist treatment. The core insight is that paradoxes mark the limit of semantic generality, and bicontextualism is designed to track that limit without abandoning the possibility of absolute generality altogether.

BIBLIOGRAPHY

[1] Peter Aczel. "An Introduction to Inductive Definitions." In: *Handbook of Mathematical Logic*. Ed. by Jon Barwise. Vol. 90. Studies in Logic and the Foundations of Mathematics. Elsevier, 1977, pp. 739–782. DOI: https://doi.org/10.1016/S0049-237X(08)71120-0.

[2] Andrew Bacon. *A Philosophical Introduction to Higher-order Logics*. Routledge, Chapman & Hall, Incorporated, 2023.

[3] Andrew Bacon. "Mathematical Modality: An Investigation in Higher-order Logic." In: *Journal of Philosophical Logic* 53.1 (2024), pp. 131–179. DOI: 10.1007/s10992-023-09728-1.

[4] Guillermo Badia, Zach Weber, and Patrick Girard. "Paraconsistent metatheory: new proofs with old tools." In: *Journal of Philosophical Logic* 51.4 (2022), pp. 825–856. DOI: https://doi.org/10.1007/s10992-022-09651-x.

[5] Timothy Bays. "Skolem's Paradox." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Fall 2022. Metaphysics Research Lab, Stanford University, 2022. URL: https://plato.stanford.edu/archives/fall2022/entries/paradox-skolem/.

[6] Jeffrey C. Beall. "Off-Topic: A New Interpretation of Weak-Kleene Logic." In: *Australasian Journal of Logic* 13.6 (2016). DOI: 10.26686/ajl.v13i6.3976.

[7] Jeffrey C. Beall, Michael Glanzberg, and David Ripley. *Formal Theories of Truth*. First edition. Oxford: Oxford University Press, 2018.

[8] Jeffrey C. Beall and Bas C. Van Fraassen. *Possibilities and Paradox. an introduction to modal and many-valued logic*. 1. publ. Oxford [u.a.]: Oxford University Press, 2003.

[9] John L. Bell. *Higher-Order Logic and Type Theory*. Elements in Philosophy and Logic. Cambridge University Press, 2022. DOI: 10.1017/9781108981804.

[10] Sharon Berry. *A Logical Foundation for Potentialist Set Theory*. Cambridge University Press, 2022. DOI: 10.1017/9781108992756.

[11] D.A. Bochvar and Merrie Bergmann. "On a three-valued logical calculus and its application to the analysis of the paradoxes of the classical extended functional calculus." In: *History and Philosophy of Logic* 2.1-2 (1981), pp. 87–112. DOI: 10.1080/01445348108837023.

[12] George Boolos. "Nominalist Platonism." In: *The Philosophical Review* 94.3 (1985), pp. 327–344. URL: http://www.jstor.org/stable/2185003.

[13] George Boolos. "On Second-Order Logic." In: *The Journal of philosophy* 72.16 (1975), pp. 509–527.

[14] George Boolos. "Reply to Charles Parsons' "Sets and Classes"." In: *Logic, Logic, and Logic*. Ed. by Richard Jeffrey. Harvard University Press, 1998, pp. 30–36.

[15] George Boolos. "The Iterative Conception of Set." In: *The Journal of Philosophy* 68.8 (1971), pp. 215–231. URL: http://www.jstor.org/stable/2025204.

[16] George Boolos. "To Be is to be a Value of a Variable (or to be Some Values of Some Variables)." In: *The Journal of Philosophy* 81.8 (1984), pp. 430–449. URL: http://www.jstor.org/stable/2026308.

[17] George Boolos, John P. Burgess, and Richard C. Jeffrey. *Computability and Logic*. 5th ed. Cambridge University Press, 2007. DOI: 10.1017/CBO9780511804076.

[18] Tim Button and Sean Walsh. *Philosophy and model theory*. First edition. Oxford: Oxford University Press, 2018.

[19] Andrea Cantini. "Notes on Formal Theories of Truth." In: *Mathematical Logic Quarterly* 35.2 (1989), pp. 97–130. DOI: https://doi.org/10.1002/malq.19890350202.

[20] Georg Cantor. *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*. Repr. d. Ausg. Berlin 1932. Berlin ; Heidelberg [u.a.]: Springer, 1980, VII, 489 S.

[21] Richard L. Cartwright. "Speaking of Everything." In: *Noûs* 28.1 (1994), pp. 1–20. DOI: 10.2307/2215917.

[22] Cezary Cieśliński. *The Epistemic Lightness of Truth: Deflationism and its Logic*. Cambridge University Press, 2017.

[23] Sam Cowling. "Ideology and Ontology." In: *The Routledge Handbook of Metametaphysics*. Ed. by J.T.M. Miller Ricki Bliss. Taylor and Francis, 2020, pp. 376–386. DOI: 10.4324/9781315112596.

[24] Laura Crosilla and Øystein Linnebo. "Definiteness in Early Set Theory." In: *Journal for the Philosophy of Mathematics* 1 (2024), pp. 43–62. DOI: 10.36253/jpm-2933.

[25] Laura Crosilla and Øystein Linnebo. "Weyl and two kinds of potential domains." In: *Noûs* 58.2 (2024), pp. 409–430. DOI: https://doi.org/10.1111/nous.12457.

[26] Dirk van Dalen. *Logic and structure*. 5. ed. Universitext. Berlin ; Heidelberg [u.a.]: Springer, 2012.

[27] Keith J. Devlin. *Constructibility*. eng. Perspectives in mathematical logic. Berlin ; Heidelberg [u.a.]: Springer, 1984, XI, 425 S. DOI: https://doi.org/10.1017/9781316717219.

[28] Max Dickmann. *Large infinitary languages. model theory*. Amsterdam: North-Holland, 1975.

[29] Michael Dummett. *Frege: Philosophy of Language*. Harvard University Press, 1981.

[30] Michael Dummett. *Frege: Philosophy of Mathematics*. Duckworth, 1991. ISBN: 9780715626603.

[31] Michael Dummett. *The Seas of Language*. Oxford University Press, 1996. DOI: 10.1093/0198236212.001.0001.

[32] Heinz-Dieter Ebbinghaus, Jörg Flum, and Wolfgang Thomas. *Mathematical logic*. Third edition. Graduate texts in mathematics. Cham: Springer, 2021. DOI: 10.1007/978-3-030-73839-6.

[33] Solomon Feferman. "Reflecting on Incompleteness." In: *The Journal of Symbolic Logic* 56.1 (1991), pp. 1–49. URL: http://www.jstor.org/stable/2274902.

[34] Hartry Field. *Saving truth from paradox*. eng. 1. publ. Oxford: Oxford University Press, 2008.

[35] Kit Fine. "Our Knowledge of Mathematical Objects." In: *Oxford Studies In Epistemology*. Oxford University Press, 2005. DOI: 10.1093/oso/9780199285891.003.0004.

[36] Kit Fine. "Relatively Unrestricted Quantification." In: *Absolute Generality*. Ed. by Augustín Rayo and Gabriel Uzquiano. Oxford University Press, 2006, pp. 20–44. DOI: 10.1093/oso/9780199276424.003.0002.

[37] Kit Fine. "Response to Alan Weir." In: *Dialectica* 61.1 (2007), pp. 117–125. URL: http://www.jstor.org/stable/42970941.

[38] Salvatore Florio. "Unrestricted Quantification." In: *Philosophy Compass* 9.7 (2014), pp. 441–454. DOI: 10.1111/phc3.12127.

[39] Salvatore Florio and Nicholas K. Jones. "Two conceptions of absolute generality." In: *Philosophical studies* 180.5-6 (2023), pp. 1601–1621. DOI: 10.1007/s11098-022-01908-0.

[40] Salvatore Florio and Nicholas K. Jones. "Unrestricted Quantification and the Structure of Type Theory." In: *Philosophy and Phenomenological Research* 102.1 (2021), pp. 44–64. DOI: 10.1111/phpr.12621.

[41] Salvatore Florio and Øystein Linnebo. "Critical Plural Logic." In: *Philosophia Mathematica* 28.2 (2020), pp. 172–203. DOI: 10.1093/philmat/nkaa020.

[42] Salvatore Florio and Øystein Linnebo. "On the Innocence and Determinacy of Plural Quantification." In: *Noûs* 50.3 (2016), pp. 565–583. DOI: https://doi.org/10.1111/nous.12091. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/nous.12091.

[43] Salvatore Florio and Øystein Linnebo. *The Many and the One: A Philosophical Study of Plural Logic*. Oxford: Oxford University Press, 2021.

[44] Kentaro Fujimoto. "Deflationism beyond arithmetic." In: *Synthese* 196.3 (2019), pp. 1045–1069. URL: http://www.jstor.org/stable/45096389 (visited on 09/11/2024).

[45] Kentaro Fujimoto. "Predicativism About Classes." In: *Journal of Philosophy* 116.4 (2019), pp. 206–229. DOI: 10.5840/jphil2019116413.

[46] Christopher Gauker. "Against Stepping Back: A Critique of Contextualist Approaches to the Semantic Paradoxes." In: *Journal of Philosophical Logic* 35.4 (2006), pp. 393–422. DOI: 10.1007/s10992-006-9026-y.

[47] Michael Glanzberg. "A Contextual-Hierarchical Approach to Truth and the Liar Paradox." In: *Journal of Philosophical Logic* 33.1 (2004), pp. 27–88. URL: http://www.jstor.org/stable/30226945.

[48] Michael Glanzberg. "Context and Unrestricted Quantification." In: *Absolute Generality*. Ed. by Augustín Rayo and Gabriel Uzquiano. Oxford University Press, 2006, pp. 45–74. DOI: 10.1093/oso/9780199276424.003.0003.

[49] Michael Glanzberg. "Quantification and Realism." In: *Philosophy and Phenomenological Research* 69.3 (2004), pp. 541–572. URL: http://www.jstor.org/stable/40040767.

[50] Michael Glanzberg. "The Liar in Context." In: *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 103.3 (2001), pp. 217–251. URL: http://www.jstor.org/stable/4321137.

[51] Michael Glanzberg. "Truth, Reflection, and Hierarchies." In: *Synthese* 142.3 (2004), pp. 289–315. URL: http://www.jstor.org/stable/20118515.

[52] Kurt Gödel. "What is Cantor's Continuum Problem?" In: *The American Mathematical Monthly* 54.9 (1947), pp. 515–525. URL: http://www.jstor.org/stable/2304666.

[53] Kurt Gödel. "What is Cantor's continuum problem?" In: *Philosophy of Mathematics: Selected Readings*. Ed. by Paul Benacerraf and HilaryEditors Putnam. 2nd ed. Revised and expanded version of Gödel [52]. Cambridge University Press, 1984, pp. 470–485. DOI: 10.1017/CBO9781139171519.025.

[54] Owen Griffiths and A.C. Paseau. *One True Logic: A Monist Manifesto*. Oxford University Press, 2022. DOI: 10.1093/oso/9780198829713.001.0001.

[55] Berta Grimau. "In defence of Higher-Level Plural Logic: drawing conclusions from natural language." In: *Synthese* 198.6 (2021), pp. 5253–5280. DOI: 10.1007/s11229-019-02399-z.

[56] Petr Hájek and Pavel Pudlák. *Metamathematics of first-order arithmetic*. 2nd printing 1998 of the 1st edition 1993. Perspectives in mathematical logic. Berlin ; Heidelberg [u.a.]: Springer, 1998.

[57] Volker Halbach. *Axiomatic theories of truth*. Cambridge: Cambridge University Press, 2011.

[58] Volker Halbach and Leon Horsten. "Axiomatizing Kripke's Theory of Truth." In: *Journal of Symbolic Logic* 71.2 (2006), pp. 677–712. DOI: 10.2178/jsl/1146620166.

[59] Bob Hale and Crispin Wright. "Logicism in the Twenty-First Century." In: *Oxford Handbook of Philosophy of Mathematics and Logic*. Ed. by Stewart Shapiro. Oxford University Press, 2005. DOI: 10.1093/0195148770.003.0006.

[60] Bob Hale and Crispin Wright. *The Reason's Proper Study: Essays Towards a Neo-Fregean Philosophy of Mathematics*. Oxford scholarship online. Clarendon Press, 2001.

[61] Joel David Hamkins. "The Set-Theoretic Multiverse." In: *The Review of Symbolic Logic* 5.3 (2012), pp. 416–449. DOI: 10.1017/S1755020311000359.

[62] William Porter Hanf and Dana Scott. "Classifying inaccessible cardinals." In: *Notices of the American Mathematical Society* 8 (1961), p. 445.

[63] Kai Hauser. "Cantor's Absolute in Metaphysics and Mathematics." In: *International Philosophical Quarterly* 53.2 (2013), pp. 161–188.

[64] Richard Kimberly Heck. "Self-reference and the Languages of Arithmetic." In: *Philosophia Mathematica* 15.1 (2007). (originally published under the name "Richard G. Heck, Jr"), pp. 1–29. DOI: 10.1093/philmat/nkl028.

[65] Geoffrey Hellman. *Mathematics Without Numbers: Towards a Modal-Structural Interpretation*. Oxford: Oxford University Press, 1989. DOI: 10.2307/2185960.

[66] Simon Hewitt. "When Do Some Things Form a Set?" In: *Philosophia Mathematica* 23.3 (2015), pp. 311–337. DOI: 10.1093/philmat/nkv010.

[67] James Higginbotham. "GB Theory: An Introduction." In: *Handbook of Logic and Language*. Ed. by Johan van Benthem and Alice ter Meulen. Amsterdam: North-Holland, 1997, pp. 311–360. DOI: https://doi.org/10.1016/B978-044481714-3/50008-4.

[68] James Higginbotham. "On Second-Order Logic and Natural Language." In: *Between logic and intuition: essays in honor of Charles Parsons*. Ed. by Gila Sher and Richard Tieszen. Cambridge University Press, 2000, pp. 79–99.

[69] David Hilbert. "On the infinite." In: *Philosophy of Mathematics: Selected Readings*. Ed. by Paul Benacerraf and HilaryEditors Putnam. 2nd ed. Cambridge University Press, 1984, pp. 183–201. DOI: https://doi.org/10.1017/CBO9781139171519.010.

[70] Wilfrid Hodges. *Model Theory*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1993. DOI: https://doi.org/10.1017/CBO9780511551574.

[71] Leon Horsten. *The Tarskian Turn: Deflationism and Axiomatic Truth*. The MIT Press, July 2011. DOI: 10.7551/mitpress/9780262015868.001.0001.

[72] Andrea Iacona. *Propositions*. Filosofia. Genova: Name, 2002.

[73] Luca Incurvati. *Conceptions of Set and the Foundations of Mathematics*. Cambridge University Press, 2020. DOI: 10.1017/9781108596961.

[74] Luca Incurvati. "Everything, More or Less: A Defence of Generality Relativism, by J. P. Studd." In: *Mind* 131.524 (June 2021), pp. 1311–1321. DOI: 10.1093/mind/fzaa096.

[75] Thomas Jech. *Set theory*. The third millennium edition, revised and expanded, corrected 4th printing. Springer monographs in mathematics. Berlin ; Heidelberg: Springer, 2006. DOI: https://doi.org/10.1007/3-540-44761-X.

[76] Akihiro Kanamori. "Zermelo and Set Theory." In: *The Bulletin of Symbolic Logic* 10.4 (2004), pp. 487–553. URL: http://www.jstor.org/stable/3216738.

[77] Carol Karp. *Languages with expressions of infinite length*. Studies in Logic and the Foundations of Mathematics ; v. 36. Amsterdam: North-Holland, 1964.

[78] Stephen Cole Kleene. "On Notation for Ordinal Numbers." In: *The Journal of Symbolic Logic* 3.4 (1938), pp. 150–155. URL: http://www.jstor.org/stable/2267778.

[79] Saul Kripke. "Outline of a Theory of Truth." In: *The Journal of Philosophy* 72.19 (1975), pp. 690–716. URL: http://www.jstor.org/stable/2024634.

[80] Kenneth Kunen. *Set theory. an introduction to independence proofs*. Studies in logic and the foundations of mathematics. Amsterdam: North-Holland, 1980.

[81] Shaughan Lavine. "Something About Everything: Universal Quantification in the Universal Sense of Universal Quantification." In: *Absolute Generality*. Ed. by Augustín Rayo and Gabriel Uzquiano. Oxford University Press, 2006, pp. 98–148.

[82] Graham Leigh and Carlo Nicolai. "Axiomatic Truth, Syntax and Metatheoretic Reasoning." In: *The Review of Symbolic Logic* 6.4 (2013), pp. 613–636. DOI: 10.1017/S1755020313000233.

[83] David Lewis. *Parts of classes*. Oxford: Blackwell, 1991.

[84] Øystein Linnebo. "Absolute Generality." In: *Oxford Handbook of Philosophy of Logic*. Ed. by Filippo Ferrari et al. Oxford University Press, forthcoming.

[85] Øystein Linnebo. "Dummett on Indefinite Extensibility." In: *Philosophical Issues* 28.1 (2018), pp. 196–220. DOI: https://doi.org/10.1111/phis.12122.

[86] Øystein Linnebo. "Plural Quantification." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2022. Metaphysics Research Lab, Stanford University, 2022.

[87] Øystein Linnebo. "Plural Quantification Exposed." In: *Noûs* 37.1 (2003), pp. 71–92. URL: http://www.jstor.org/stable/3506205.

[88] Øystein Linnebo. "Pluralities and Sets." In: *Journal of Philosophy* 107.3 (2010), pp. 144–164. URL: https://www.jstor.org/stable/25700491.

[89] Øystein Linnebo. "Sets, Properties, and Unrestricted Quantification." In: *Absolute Generality*. Ed. by Augustín Rayo and Gabriel Uzquiano. Oxford University Press, 2006, pp. 149–178. DOI: 10.1093/oso/9780199276424.003.0006.

[90] Øystein Linnebo. "The Potential Hierarchy of Sets." In: *The Review of Symbolic Logic* 6.2 (2013), pp. 205–228. DOI: 10.1017/S1755020313000014.

[91] Øystein Linnebo. *Thin Objects: An Abstractionist Account*. Oxford: Oxford University Press, 2018.

[92] Øystein Linnebo and David Nicolas. "Superplurals in English." In: *Analysis* 68.3 (2008), pp. 186–197. DOI: 10.1093/analys/68.3.186.

[93] Øystein Linnebo and Agustìn Rayo. "Hierarchies Ontological and Ideological." In: *Mind* 121.482 (2012), pp. 269–308. URL: https://www.jstor.org/stable/23321978.

[94] Øystein Linnebo and Stewart Shapiro. "Actual and Potential Infinity." In: *Noûs* 53.1 (2019), pp. 160–191. DOI: https://doi.org/10.1111/nous.12208.

[95] Øystein Linnebo and Stewart Shapiro. "Predicative Classes and Strict Potentialism." In: *Philosophia Mathematica* (Nov. 2024), pp. 1–37. DOI: 10.1093/philmat/nkae020.

[96] Øystein Linnebo and Stewart Shapiro. "Predicativism as a Form of Potentialism." In: *The Review of Symbolic Logic* 16.1 (2023), pp. 1–32. DOI: 10.1017/S1755020321000423.

[97] Penelope Maddy. "Believing the Axioms. I." In: *The Journal of Symbolic Logic* 53.2 (1988), pp. 481–511. URL: http://www.jstor.org/stable/2274520.

[98] Penelope Maddy and Jouko Väänänen. *Philosophical Uses of Categoricity Arguments*. Elements in the Philosophy of Mathematics. Cambridge University Press, 2023.

[99] Poppy Mankowitz. "Paradox and Context Shift." In: *Philosophical Studies* 180.5-6 (2022), pp. 1539–1557. DOI: 10.1007/s11098-022-01841-2.

[100] Vann McGee. "Everything." In: *Between Logic and Intuition: Essays in Honor of Charles Parsons*. Ed. by Gila Sher and RichardEditors Tieszen. Cambridge: Cambridge University Press, 2000, pp. 54–78.

[101] Vann McGee. "There's a Rule for Everything." In: *Absolute Generality*. Ed. by Augustín Rayo and Gabriel Uzquiano. Oxford: Oxford University Press, 2006, pp. 75–97. DOI: https://doi.org/10.1093/oso/9780199276424.003.0007.

[102] Vann McGee. *Truth, Vagueness, and Paradox. an essay on the logic of truth*. Indianapolis: Hackett, 1991.

[103] Toby Meadows. "Sets and supersets." In: *Synthese (Dordrecht)* 193.6 (2016), pp. 1875–1907. DOI: 10.1007/s11229-015-0818-x.

[104] Richard Montague. "Reductions of Higher-Order Logic." In: *The Theory of Models*. Ed. by J.W. Addison, Leon Henkin, and Alfred Tarski. Studies in Logic and the Foundations of Mathematics. North-Holland, 2014, pp. 251–264. ISBN: 978-0-7204-2233-7.

[105] Sean Morris. *Quine, New Foundations, and the Philosophy of Set Theory*. Cambridge University Press, 2018.

[106] Yiannis Nicholas Moschovakis. *Elementary Induction on Abstract Structures*. Mineola, N.Y.: Dover Publications, 1974.

[107] Andrzej Mostowski. "An undecidable arithmetical statement." In: *Fundamenta mathematicae* 36 (1949), pp. 143–164. DOI: 10.4064/fm-36-1-143-164.

[108] Beau Madison Mount and Daniel Waxman. "Stable and Unstable Theories of Truth and Syntax." In: *Mind* 130.518 (2021), pp. 439–473. DOI: 10.1093/mind/fzaa034.

[109] Julien Murzi and Lorenzo Rossi. "Reflection Principles and the Liar in Context." In: *Philosophers' Imprint* 18 (2018).

[110] Julien Murzi and Lorenzo Rossi. *Truth and Paradox in Context*. Oxford, in preparation.

[111] John von Neumann. "Über eine Widerspruchfreiheitsfrage in der axiomatischen Mengenlehre." In: *Journal für die reine und angewandte Mathematik* 160 (1929), pp. 227–241. URL: http://eudml.org/doc/149688.

[112] Carlo Nicolai. "A Note on Typed Truth and Consistency Assertions." In: *Journal of Philosophical Logic* 45.1 (2016), pp. 89–119. URL: http://www.jstor.org/stable/43895429.

[113] Carlo Nicolai. "Deflationary Truth and the Ontology of Expressions." In: *Synthese* 192 (2015), pp. 4031–4055. DOI: https://doi.org/10.1007/s11229-015-0729-x.

[114] Alex Oliver and Timothy Smiley. *Plural Logic: Second Edition, Revised and Enlarged*. Oxford: Oxford University Press, 2016.

[115] Charles Parsons. "Frege's Theory of Numbers." In: *Philosophy in America*. Ed. by Max Black. 1964, pp. 180–203.

[116] Charles Parsons. "Sets and Classes." In: *Noûs* 8.1 (1974), pp. 1–12. URL: http://www.jstor.org/stable/2214641.

[117] Charles Parsons. "Sets and Modality." In: *Mathematics in philosophy: selected essays*. Ed. by Charles Parsons. Cornell University Press, 1983, pp. 298–341.

[118] Charles Parsons. "The Liar Paradox." In: *Journal of Philosophical Logic* 3.4 (1974), pp. 381–412. URL: http://www.jstor.org/stable/30226094.

[119] Charles Parsons. "The Problem of Absolute Universality." In: *Absolute Generality*. Ed. by Augustín Rayo and Gabriel Uzquiano. Oxford University Press, 2006, pp. 203–219. DOI: 10.1093/oso/9780199276424.003.0008.

[120] A.C. Paseau and Owen Griffiths. "Is English consequence compact?" In: *Thought (Hoboken, N.J.)* 10.3 (2021), pp. 188–198.

[121] Lavinia Picollo. "Reference in Arithmetic." In: *The Review of Symbolic Logic* 11.3 (2018), pp. 573–603. DOI: 10.1017/S1755020317000351.

[122] Lavinia Picollo and Thomas Schindler. "Deflationism and the Function of Truth." In: *Philosophical Perspectives* 32.1 (2018), pp. 326–351. DOI: https://doi.org/10.1111/phpe.12113.

[123] Lavinia Picollo and Thomas Schindler. "Higher-Order Logic and Disquotational Truth." In: *Journal of Philosophical Logic* 51.4 (2022), pp. 879–918. DOI: https://doi.org/10.1007/s10992-022-09654-8.

[124] Wolfram Pohlers. *Proof Theory. The first step into impredicativity*. Springer Berlin, Heidelberg, 2009. DOI: https://doi.org/10.1007/978-3-540-69319-2.

[125] Michael Potter. *Set Theory and its Philosophy: A Critical Introduction*. Oxford University Press, 2004. DOI: 10.1093/acprof:oso/9780199269730.001.0001.

[126] Graham Priest. *In Contradiction: A Study of the Transconsistent*. Oxford: Oxford University Press, 2006.

[127] Willard van Orman Quine. "New Foundations for Mathematical Logic." In: *The American Mathematical Monthly* 44.2 (1937), pp. 70–80. URL: http://www.jstor.org/stable/2300564.

[128] Willard van Orman Quine. "On What There Is." In: *From a Logical Point of View*. Ed. by Willard Van Orman Quine. Harvard University Press, 1953, pp. 1–19.

[129] Willard van Orman Quine. "Ontology and Ideology." In: *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 2.1 (1951), pp. 11–15. URL: http://www.jstor.org/stable/4318102.

[130] Willard van Orman Quine. *Philosophy of logic*. 2. ed. Cambridge: Harvard University Press, 1986.

[131] Michael Rathjen. "The Realm of Orinal Analysis." In: *Sets and Proofs*. Ed. by S. Barry Cooper and John K Truss. London Mathematical Society Lecture Note Series. Cambridge University Press, 1999, pp. 219–280.

[132] Agustìn Rayo. "Beyond Plurals." In: *Absolute Generality*. Ed. by Augustín Rayo and Gabriel Uzquiano. Oxford University Press, 2006, pp. 220–254. DOI: 10.1093/oso/9780199276424.003.0009.

[133] Agustin Rayo and Stephen Yablo. "Nominalism Through De-Nominalization." In: *Noûs* 35.1 (2001), pp. 74–92. URL: http://www.jstor.org/stable/2671946 (visited on 01/21/2024).

[134] Agustìn Rayo and Gabriel Uzquiano. *Absolute Generality*. Ed. by Augustín Rayo and Gabriel Uzquiano. Oxford University Press, 2006. DOI: 10.1093/oso/9780199276424.001.0001.

[135] Agustìn Rayo and Gabriel Uzquiano. "Toward a Theory of Second-Order Consequence." In: *Notre Dame Journal of Formal Logic* 40.3 (1999), pp. 315–325. DOI: 10.1305/ndjfl/1022615612.

[136] Agustìn Rayo and Timothy Williamson. "A Completeness Theorem for Unrestricted First-Order Languages." In: *Liars and Heaps – New Essays on Paradox*. Ed. by JC Beall. Oxford: Oxford University Press, 2003, pp. 331–356.

[137] William N. Reinhardt. "Satisfaction Definitions and Axioms of Infinity in a Theory of Properties With Necessity Operator." In: *Mathematical Logic in Latin America*. Ed. by A.I. Arruda, R. Chuaqui, and N.C.A. Da Costa. Vol. 99. Studies in Logic and the Foundations of Mathematics. Elsevier, 1980, pp. 267–303. DOI: https://doi.org/10.1016/S0049-237X(09)70491-4.

[138] William N. Reinhardt. "Some Remarks on Extending and Interpreting Theories with a Partial Predicate for Truth." In: *Journal of Philosophical Logic* 15.2 (1986), pp. 219–251. URL: http://www.jstor.org/stable/30226351.

[139] Michael D. Resnik. "Second-Order Logic Still Wild." In: *The Journal of Philosophy* 85.2 (1988), pp. 75–87. URL: http://www.jstor.org/stable/2026993.

[140] Greg Restall. "A Note on Naive Set Theory in LP." In: *Notre Dame Journal of Formal Logic* 33.3 (1992), pp. 422–432. DOI: 10.1305/ndjfl/1093634406.

[141] Lucas Rosenblatt. "Should the Non-Classical Logician be Embarrassed?" In: *Philosophy and Phenomenological Research* 104.2 (2022), pp. 388–407. DOI: https://doi.org/10.1111/phpr.12770.

[142] Marcus Rossberg. "Somehow Things Do Not Relate: On the Interpretation of Polyadic Second-Order Logic." In: *Journal of Philosophical Logic* 44.3 (2015), pp. 341–350. DOI: https://doi.org/10.1007/s10992-014-9326-6.

[143] Lorenzo Rossi. "A Unified Theory of Truth and Paradox." In: *The Review of Symbolic Logic* 12.2 (2019), pp. 209–254. DOI: 10.1017/S1755020319000078.

[144] Lorenzo Rossi. "Bicontextualism." In: *Notre Dame Journal of Formal Logic* 64.1 (2023), pp. 95–127. DOI: 10.1215/00294527-2023-0004.

[145] Bertrand Russell. "Mathematical Logic as based on the Theory of Types." In: *American Journal of Mathematics* 30 (1908), p. 222.

[146] Bertrand Russell. "On Some Difficulties in the Theory of Transfinite Numbers and Order Types." In: *Proceedings of the London Mathematical Society* s2-4.1 (1907), pp. 29–53. DOI: https://doi.org/10.1112/plms/s2-4.1.29.

[147] Thomas Schindler. "Classes, why and how." In: *Philosophical Studies* 176.2 (2019), pp. 407–435. DOI: https://doi.org/10.1007/s11098-017-1022-2.

[148] Stewart Shapiro. "All sets great and small: and I do mean ALL." In: *Philosophical Perspectives* 17.1 (2003), pp. 467–490. DOI: 10.1111/j.1520-8583.2003.00014.x.

[149] Stewart Shapiro. *Foundations without Foundationalism. A case for second-order logic*. Oxford: Oxford University Press, 1991. DOI: 10.1093/0198250290.001.0001.

[150] Keith Simmons. *Universality and the Liar: An Essay on Truth and the Diagonal Argument*. Cambridge University Press, 1993. DOI: 10.1017/CBO9780511551499.

[151] Daniel Stoljar. "Physicalism." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Spring 2024. Metaphysics Research Lab, Stanford University, 2024.

[152] James Studd. *Everything, more or less. a defence of generality relativism*. First edition. Oxford scholarship online. Oxford: Oxford University Press, 2019. DOI: 10.1093/oso/9780198719649.001.0001.

[153] James Studd. "The Iterative Conception of Set: A (Bi-) Modal Axiomatisation." In: *Journal of Philosophical Logic* 42.5 (2013), pp. 697–725. URL: http://www.jstor.org/stable/42001254.

[154] James Studd. "V – Generality, Extensibility, and Paradox." In: *Proceedings of the Aristotelian Society* 117.1 (2017), pp. 81–101. DOI: 10.1093/arisoc/aox003.

[155] William Tait. "Finitism." In: *Journal of Philosophy* 78.9 (1981), pp. 524–546. DOI: 10.2307/2026089.

[156] Alfred Tarski. "Der Wahrheitsbegriff in den Formalisierten Sprachen." In: *Studia Philosophica* 1 (1935), pp. 261–405.

[157] Neil Tennant. "Logicism and Neologicism." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Winter 2023. Metaphysics Research Lab, Stanford University, 2023.

[158] Robert Trueman. *Properties and Propositions: The Metaphysics of Higher-Order Logic*. Cambridge University Press, 2021.

[159] Gabriel Uzquiano. "Plural Quantification and Classes." In: *Philosophia Mathematica* 11.1 (2003), pp. 67–81. DOI: 10.1093/philmat/11.1.67.

[160] Gabriel Uzquiano. "Varieties of Indefinite Extensibility." In: *Notre Dame Journal of Formal Logic* 56.1 (2015), pp. 147–166. DOI: 10.1215/00294527-2835056.

[161] Jouko Väänänen. "Second-order and Higher-order Logic." In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2021. Metaphysics Research Lab, Stanford University, 2021.

[162] Albert Visser. "Four valued semantics and the Liar." In: *Journal of Philosophical Logic* 13.2 (1984), pp. 181–212. DOI: 10.1007/BF00453021.

[163] Hao Wang. *A Logical Journey: From Gödel to Philosophy*. The MIT Press, 1997. ISBN: 9780262285766. DOI: 10.7551/mitpress/4321.001.0001. URL: https://doi.org/10.7551/mitpress/4321.001.0001.

[164] Zach Weber. *Paraconsistency in Mathematics*. Elements in the Philosophy of Mathematics. Cambridge University Press, 2022. DOI: https://doi.org/10.1017/9781108993968.

[165] Zach Weber. *Paradoxes and Inconsistent Mathematics*. Cambridge University Press, 2021. DOI: https://doi.org/10.1017/9781108993135.

[166] Alan Weir. "Is it too much to Ask, to Ask for Everything?" In: *Absolute Generality*. Ed. by Augustín Rayo and Gabriel Uzquiano. Oxford University Press, 2006, pp. 333–368. DOI: 10.1093/oso/9780199276424.003.0012.

[167] Alan Weir. "Naive set theory is innocent!" In: *Mind* 107.428 (1998), pp. 763–798. DOI: 10.1093/mind/107.428.763.

[168] Philip Welch and Leon Horsten. "Reflecting on Absolute Infinity." In: *The Journal of Philosophy* 113.2 (2016), pp. 89–111.

[169] Timothy Williamson. "Everything." In: *Philosophical Perspectives* 17 (2003), pp. 415–465. URL: http://www.jstor.org/stable/3840881.

[170] Timothy Williamson. "Indefinite Extensibility." In: *Grazer philosophische Studien* 55.1 (1999), pp. 1–24. DOI: 10.5840/gps19985512.

[171] Peter W. Woodruff. "Paradox, Truth and Logic Part I: Paradox and Truth." In: *Journal of Philosophical Logic* 13.2 (1984), pp. 213–232. URL: http://www.jstor.org/stable/30227029.

[172] Crispin Wright. *Frege's Conception of Numbers as Objects*. Aberdeen: Aberdeen University Press, 1983.

[173] Ernst Zermelo. "Über Grenzzahlen Und Mengenbereiche: Neue Untersuchungen Über Die Grundlagen der Mengenlehre." In: *Ernst Zermelo, Collected Works, Gesammelte Werke, Volume I*. Ed. by Heinz-Dieter Ebbinghaus, Craig G. Fraser, and Akihiro Kanamori. Springer, 2010, pp. 400–429.

DECLARATION

I hereby declare that all the work presented in this thesis is my own unless indicated otherwise.

Munich, June 29, 2025

Simon Schmitt