Properties Grounded in Identity

A Study Of Essential Properties

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Philosophie an der

Ludwig-Maximilians-Universität
München

vorgelegt von

Johannes Korbmacher

aus Wesel

2016

Erstgutachter:	Prof. Dr. Dr. Hannes Leitgeb
Zweitgutachter:	Prof. Dr. Benjamin Schnieder
Datum der mündlichen Prüfung:	04.02.2016

Contents

Acknowledgements

There are a great many people that I'm indebted to more than I can possibly put into words. But I should like to thank a few of them for the important role they've played in my philosophical development and for making this dissertation possible.

Three people stand out in their influence on my philosophical development. My father Bruno Korbmacher was the first to introduce me to philosophy, and I shall always be grateful for that. In Münster, my friend and teacher Martin Pleitz was instrumental in developing and sharpening my philosophical skills. Martin has taught me the value of discussion and he has always been a worthy opponent in our philosophical quibbles. Finally, in Munich, my supervisor Hannes Leitgeb quickly became the role model of a philosopher to me. I could not have imagined a greater teacher under whose guidance to study.

Over the last four years, being a member of MCMP felt like being a kid in a candy store—a *philosophical* candy store, that is. I have met so many great philosophers—too many to count[1]—and I have learned so many things—too many to recount. Of the many great people I've met at MCMP a few stand out for their invaluable help in writing this dissertation. First of all, nobody has read more of this dissertation than Albert J. J. Anglberger—except for me of course. Without his comments this dissertation would not have been possible. Only Thomas Schindler, Johannes Stern, Lavinia Picollo, and Martin Fischer have made it possible for me to write the second chapter "Axiomatic Theories of Ground"—without their invaluable comments and patient answers to my relentless questions about dots, I could not have tackled this project. Ole Thomassen Hjortland and Norbert Gratzl have been amazing teachers and friends—they taught me much more than just how to perform a proper cut-elimination argument. The whole MCMP-gang has had a lasting influence on me: Hans-Christoph Kotzsch, Thomas Meier, Georg Schiemer, Jeffrey Ketland, Seamus Bradley, Karim Thébault, and all the

[1] Somehow, Hannes Leitgeb has managed to bring more than ω-many people to Munich—a feat not easily accomplished.

countless others—you guys have been great! From Hamburg, I would like to thank Benjamin Schnieder for his kind agreement to be the second examiner of this thesis. And I should like to thank the Alexander von Humboldt foundation for funding my PhD fellowship at the MCMP.

On a more personal note, I would like to thank my mother Birgit Hille-Korbmacher for giving me the home I needed when I needed it. My sister Elisabeth and my brother Raphael I would like to thank for reminding me of the value of family. My grandfather Hermann-Josef Hille, himself quite interested in philosophy, has always engaged with me in philosophical discussion, he's provided me with books, and even articles from newspapers. By doing so he's fostered my philosophical ambitions from an early age and I would not have become the philosopher I am today without him. And finally, I would like to thank Chiara Lisciandra. She has changed me in so many ways, as a philosopher and as a man. She has kept me grounded, healthy, and motivated—but most of all, she has kept me sane through all of this.

Johannes Korbmacher
Munich
November 2015

Deutschsprachige Zusammenfassung

Das Thema dieser Dissertation ist *Essentialismus*: die These, dass zumindest einige Dinge zumindest einige essentielle Eigenschaften haben. Das Ziel der Dissertation ist es, eine *Explikation* (im Sinne Carnaps) des Begriffes essentieller Eigenschaften zu geben, wobei der Autor den Essentialismus als gegeben voraussetzt.

Das erste Kapitel "Introduction" stellt die Einleitung der Dissertation dar. In diesem Kapitel gibt der Autor den notwendigen philosophischen Hintergrund und eine Motivation für das Projekt der Dissertation. Zu diesem Zweck stellt der Autor zuerst die gängige Analyse von essentiellen Eigenschaften mithilfe des Begriffs von Notwendigkeit *de re* dar und erläutert die notwendigen philosophischen Annahmen. Darauf aufbauend diskutiert der Autor kurz, wie der Begriff von Notwendigkeit *de re* durch Kripkes mögliche-Welten-Semantik gegen Quines Einwände behauptet wurde. Anschließend diskutiert der Autor Fines Einwände gegen die Analyse von essentiellen Eigenschaft durch Notwendigkeit *de re* und stellt Fines Gegenvorschlag vor, essentielle Eigenschaften als Eigenschaften zu verstehen, die in der Identität ihrer Träger begründet liegen. Der Autor argumentiert dafür, das Problem der Analyse des Begriffs von essentiellen Eigenschaften als ein *Explikationsproblem* im Sinne Carnaps zu verstehen. Dies bedeutet insbesondere, dass der Autor annimmt, dass es keine "richtige" oder "falsche" Analyse des Begriffes gibt, sondern die Standards, an denen ein Analysevorschlag gemessen werden sollte, in der *Nützlichkeit* für die intendierten Anwendungen der Analyse begründet liegen. Aufbauend darauf argumentiert der Autor dafür, dass es mindestens *zwei* Begriffe von essentiellen Eigenschaften gibt: einen, der durch die Analyse mithilfe von Notwendigkeit *de re* expliziert wird, und einen, der durch Fines Analyse mithilfe von metaphysischer Begründung expliziert wird. Allerdings attestiert der Autor, dass Fines Analyse, zumindest dem gegenwärtigen Stand der Forschung nach, nicht vergleichbar präzise wie die Analyse mithilfe von Notwendigkeit *de re* formuliert ist. Insbesondere, so der Autor, fehlen zwei wesentliche Aspekte:

(1) eine vollständige (semantische) Analyse metaphysischer Begründung, vergleichbar mit der mögliche-Welten-Semantik von Kripke, und

(2) eine metaphysisch robuste Hintergrundtheorie von Eigenschaften, welche den Ansprüchen der Fineschen Analyse genügt.

Der Autor formuliert es als Ziel der Dissertation, diese zwei fehlenden Aspekte zu liefern.

Im zweiten Kapitel "Axiomatic Theories of Ground" verfolgt der Autor einen axiomatischen Ansatz zu Theorien von metaphysischer Begründung. Der Autor argumentiert dafür, dass ein solcher Ansatz nicht nur natürlich, sondern insbesondere auch technisch und philosophisch gut motiviert ist. Ausgehend von den üblicherweise akzeptierten Prinzipien für metaphysische Begrüngung, entwickelt der Autor eine axiomatische Theorie und zeigt, dass diese nicht nur konsistent ist, sondern insbesondere auch beweistheoretisch konservativ über der Theorie von positiver Wahrheit ist. Anschließend diskutiert der Autor eine mögliche Erweiterung seiner Theorie um getypte Wahrheitsprädikate und zeigt, dass auch diese Erweiterung konsistent bis zum einem Typisierungslevel von ϵ_0 ist. Der Autor argumentiert dafür, dass diese Theorie eine mögliche Lösung für Fines *Puzzle of Ground* darstellt. Abschließend weißt der Autor auf die technischen Grenzen seiner axiomatischen Herangehensweise hin und argumentiert dafür, dass zumindest zum gegenwärtigen Stand der Forschung, axiomatische Theorien von metaphysischer Begründung noch nicht weit genug entwickelt sind, um die Aufgabe (1) aus der Einleitung hinreichend präzise zu beantworten.

Im dritten Kapitel "The Full Logic of Worldly Ground" wendet sich der Autor operationellen Ansẗzen zu metaphysischer Begründung zu. Ziel des Kapitels ist es, die bestehenden operationalen Ansätze metaphysischer Begründung in der Hinsicht zu erweitern, dass *iterierte* Anwendungen von Begründungsoperatoren beschrieben werden können. Der Autor motiviert die Fragestellung aus philosophischen Positionen zur Begründung von metaphysischer Begründung. Der Autor argumentiert dafür, dass unterschiedliche Ansichten zur Begründung von Begründung zu unterschiedlichen Logiken metaphyischer Begründung führen. In diesem Kapitel setzt sich der Autor das Ziel, einen operationalen Ansatz, unter der These, dass metaphysische Gründe Begründung begründen, zu entwickeln. Zu diesem Zweck erweitert der Autor die Wahrmacher Semantik von Fine [42] um Wahrheitsbedingungen für iterierte Anwendungen von metaphysischen Begründungsoperatoren in infinitären Sprachen. Die Semantik soll dabei intuitiv der These, dass Gründe Begründung begründen, entsprechen. Anschließend gibt der Autor ein infinitäres Beweissystem für die Semantik und zeigt die Korrekheit des Systems bezüglich der zuvor vorgestellten Semantik. Der Autor diskutiert abschließend Möglichkeiten, die Resultate des Kapitels zu erweitern. Insbesondere weist der Autor auf Komplikationen für einen möglichen

Vollständigkeitsbeweis hin.

Im vierten Kapitel der Arbeit mit dem Titel "How to Distinguish Necessarily Equivalent But Distinct Properties" entwickelt der Autor eine neue Theorie hyperintensionaler Eigenschaften. Der Autor beginnt mit einer Kritik an den meist vertretenen hyperintensionalen Eigenschaftstheorien, insbeondere der unmögliche-Welten Theorie und der Theorie von strukturierten Eigenschaften. Anschließend entwickelt der Autor seine eigene hyperintensionale Theorie von Eigenschaften. Diese Theorie baut auf der Intuition auf, dass wir Eigenschaften über ihre *Exemplifikationskriterien* individuieren sollten. Dabei sind die Exemplifikationskriterien einer Eigenschaft für einen Gegenstand diejenigen Sachverhalte bezüglich des Gegenstands, welche der Fall sein müssen, damit der Gegenstand die Eigenschaft exemplifiziert. Der Autor argumentiert dafür, dass diese Idee zu einer natürlichen und philosophisch adäquaten Eigenschaftstheorie führt und deutet auf verschiedene Anwendungsmöglichkeiten hin.

Im Schlussteil "Conclusion" fasst der Autor die Ergebnisse der Arbeit kurz zusammen und wendet sie nocheinmal auf die konkrete Fragestellung der Dissertation an. Insbesondere stellt der Autor hier das zentrale Ergebnis der Arbeit vor: er gibt eine Explikation der Auffassung von essentiellen Eigenschaften, verstanden als Eigenschaften, die in der Identität ihrer Träger begründet liegen. Diese Explikation besagt im Wesentlichen, dass eine Eigenschaft genau dann essentiell für einen Gegenstand ist, wenn die Behauptung, dass der Gegenstand diese Eigenschaft exemplifiziert, von dem augezeichneten Identitätssachverhalt des Gegenstandes verifiziert wird. Der Autor argumentiert dafür, dass diese Analyse zusammen mit der Eigenschaftstheorie aus dem vierten Kapitel eine plausible Explikation des Begriffes von essentiellen Eigenschaften, verstanden als Eigenschaften, die in der Identität ihrer Träger begründet liegen, gibt. Abschließend weist der Autor noch auf verschiedene Möglichkeiten hin, die Ergebnisse der Arbeit weiterzuentwickeln.

Chapter 1

Introduction

The essence of a thing is what it is said to be in respect of itself.
Aristotle, *Metaphysics*, 1029b14

This dissertation is about *essentialism*: the view that at least some things have at least some *essential properties*.[1,2] To informally illustrate essentialism and the concept of essential properties, think of Socrates as an example. Typically, essentialists would say that being a man is an essential property of Socrates.[3] In contrast, essentialists would typically say that being married to Xanthippe is *not* an essential property of Socrates.[4] The aim of this dissertation is to give a *Carnapian explication* of the concept of essential properties [17, p. 1–18], while assuming essentialism. In this introduction, we'll[5] give some informal philosophical background and motivation.

[1] In this dissertation, for reasons of perspicuity, we shall often confine ourselves to properties, while leaving relations out of the picture. Properties are, of course, a special case of relations: a property is simply a unary relation. Everything that we'll say about properties in the following can easily be generalized to arbitrary relations.

[2] In this dissertation, we shall exclusively focus on what Correia [23] calls *objectual essence*: the essence or rather essential properties *of objects*. Correia contrasts objectual essence with *generic essence*: the essence or perhaps essential properties *of generic things*, including concepts, properties, and so on. Generic essence is an exciting concept of essence, which we can approach with similar methods as the ones we'll develop in this dissertation. However, for reasons of perspicuity, we'll focus exclusively on objectual essence here. We hope to extend the results of this dissertation to generic essence at a later time.

[3] For a disagreeing argument, see [134].

[4] Gerundives, like "being a man" and "being married to Xanthippe," are *canonical property designators*: expressions we normally use to denote to properties [125]. Later, we shall discuss different metaphysical theories of properties, but for now we leave the concept unanalyzed.

[5] In this dissertation we will follow the suggestion of the *Chicago Manual of Style* to use contractions whenever they increase legibility.

1.1 Essentialism and Modal Metaphysics

Let's begin with the standard analysis of essential properties for most of the twentieth century. For quite some time, essentialism and essential properties had a difficult standing in analytic metaphysics. In the middle of the twentieth century, Quine still spoke of "the metaphysical jungle of Aristotelian essentialism" [116, p. 174]. Quine's main worry with essentialism lied with the concept of necessity that is involved in the concept of essential properties. Traditionally, essentialists analyze essential properties in terms of necessity [109, 90, 110, 76, 83, 137]. One way of spelling out this analysis is in terms of Kripke's notion of *weak necessity*, i.e. necessity contingent upon existence [75, p. 138]:

Modal Analysis (MA). For all properties Φ and all objects x,

Φ is an essential property of x iff $\Box$(if x exists, then x exemplifies Φ).[6]

Here the modal operator $\Box$ expresses what philosophers typically call *metaphysical necessity*: necessity that obtains in virtue of the laws of metaphysics [49].[7] The principle **MA** captures the essentialist intuition that essential properties are simply properties that are weakly necessary to their bearers.

Given some fairly standard assumptions in philosophical modal logic, **MA** has some interesting consequences. First, we may standardly assume that whatever is necessarily the case is also actually the case. In symbols, for all statements φ, we may assume:

$$(\mathbf{T}) \quad \text{if } \Box\varphi, \text{ then } \varphi.$$

Thus, we can infer from **MA** that if a property Φ is an essential property of an object x, then if x exists, x exemplifies Φ. In other words, existing objects exemplify all their essential properties. Second, we may standardly assume that necessity and possibility are interdefinable: something is possibly the case iff it is not necessarily not the case. In symbols, for all statements φ,

[6] The term "exemplification" is a term of art from metaphysics that describes the relation that holds between an object and a property iff we would say in natural language that the object *has* the property. Thus, in metaphysical contexts, we say that an object *exemplifies* a property iff we would say in natural language that the object has the property. There are different ways, both in natural language and in metaphysics, to say that an object exemplifies a property. Take Socrates and the property of being a man as an example. We can, for example, say that Socrates *has* the property of being a man, that Socrates *bears* the property of being a man, or simply that Socrates *is* a man. We will use different locutions on different occasions, without meaning anything deep by it. Other than the previous natural language examples, we shall leave the relation of exemplification unanalyzed for now, at least until we properly discuss metaphysical theories of properties.

[7] In this introduction, we'll use formal symbols, like $\Box$, together with informal natural language expressions, like "x exists," in an attempt at *informal rigor*: the paradigm that we should be as precise as possible, even in informal exposition.

we may assume:

$$(\Diamond/\Box) \quad \Diamond\varphi \text{ iff } \neg\Box\neg\varphi,$$

where $\Diamond$ expresses (metaphysical) possibility. Thus, for an object x and a property Φ, we can infer from **MA** that if $\Diamond$(x exists & $\neg$(x exemplifies Φ)), then Φ is *not* an essential property of x.[8] In other words, properties that are *contingent* to an object—properties such that the object can exist without exemplifying them—are never essential properties of the object.

To illustrate **MA**, let's consider Socrates again. We've said that being a man is a typical example of an essential property of Socrates. By **MA** this means that $\Box$(if Socrates exists, then Socrates is a man). This is indeed plausible, for Socrates could intuitively not exist without being a man.[9] In contrast, it is intuitively plausible that Socrates could have existed without being married to Xanthippe, since, for example, the two could have never met. Thus, $\Diamond$(Socrates exists & $\neg$(Socrates is married to Xanthippe)), and so, by the previous observation, being married to Xanthippe is not an essential property of Socrates—exactly as we wish to say as essentialists. To sum up, in the cases of Socrates being a man and Socrates being married to Xanthippe, **MA** agrees with our essentialist intuitions.

It is worth pointing out one crucial background assumption here. Note that we have used Kripke's notion of *weak* necessity to spell out the modal analysis of essential properties: we've said that a property Φ is an essential property of an object x iff $\Box$(*if x exists, then* x exemplifies Φ). Now, weak necessity is necessity contingent upon existence, but instead we could have used a notion of unconditional or *strict* necessity to analyze the connection between essential properties and necessity:

Strict Modal Analysis (SMA). For all properties Φ and all objects x,

$$\Phi \text{ is an essential property of } x \text{ iff } \Box(x \text{ exemplifies } \Phi).$$

Here it is simply postulated that the essential properties of an object are the properties that the object necessarily exemplifies *regardless of whether the object exists or not.* This principle is typically endorsed by more "old-school" essentialists, like Parsons [109] and Marcus [90]. The reason why we chose weak necessity over strict necessity for the modal analysis is that we wish to subscribe to the view that Williamson [136] calls *contingentism*: the view that it's possible for some things to possibly not exist. In symbols, we can express the view as:

[8] For assume that $\Diamond$(x exists & $\neg$(x exemplifies Φ)). By $\Diamond/\Box$, this means that $\neg\Box\neg$(x exists & $\neg$(x exemplifies Φ)). But by classical logic, $\neg$(x exists & $\neg$(x exemplifies Φ)) is (logically) equivalent to (if x exists, then x exemplifies Φ). So, we get $\neg\Box$ (if x exists, then x exemplifies Φ) from $\Diamond$(x exists & $\neg$(x exemplifies Φ)). Then using *modus tollens* and **MA** we can infer that Φ is not an essential property of x. Q.E.D.

[9] But see Footnote 3.

Contingentism. $\Diamond$(for some object x $\Diamond\neg$(x exists)).

Williamson succinctly sums the view up in the slogan: "ontology is contingent" [136, p. 2]. To illustrate, consider Socrates again. We wish to allow for the possibility of Socrates never having been born, since, for example, his parents could have never met. Moreover, we wish to hold that if Socrates would never have been born, then he simply wouldn't have existed. Thus, we wish to say that $\Diamond\neg$(Socrates exists). But Socrates actually exists, and so there is some object x, namely Socrates, such that $\Diamond\neg$(x exists). Now we may standardly assume that whatever is actually the case is also possibly the case. In symbols, for all statements φ,

$$\text{if } \varphi, \text{ then } \Diamond\varphi.$$ [10]

Putting all of this together, we get **Contingentism**: Since there actually is some object x, namely Socrates, such that $\Diamond\neg$(x exists), it follows that $\Diamond$(for some object x $\Diamond\neg$(x exists)).

In contrast, Williamson [136] calls the view that necessarily everything necessarily exists *necessitism*. In symbols, we can express the view as:

Necessitism. $\Box$(for all objects x $\Box$(x exists)).

On this view, for example, we would say that even if Socrates would never have been born, he would still have existed. In fact, on the view, everything that exists, or possibly could have existed, necessarily exists. As Williamson sums up the view succinctly: "ontology is necessary" [136, p. 2]. Thus, **Necessitism** is simply the negation of **Contingentism**. In this dissertation, we do not wish to go further into the contingentism-necessitism debate.[11,12] We'll simply *assume* contingentism. Thus, the view that we're interested in might be called *contingentist essentialism*: the view that, assuming contingentism, at least some objects have at least some essential properties. For this specific view, we wish to provide an explication of the concept of essential properties.

Now, as contingentists, it's reasonable to assume **MA** rather than **SMA**. To see this, think of Socrates again. As we've said, we wish to assume that it's possible for Socrates not to have existed: $\Diamond\neg$(Socrates exists). But it's plausible to say that if Socrates doesn't exist, then he also doesn't exemplify

[10] This principle, of course, follows from **T** by instantiating φ with $\neg\varphi$, contraposition, double negation elimination, and $\Diamond/\Box$.

[11] [136] is a book length defense of **Necessitism**. Note that Williamson does *not* argue that **Contingentism** is *incoherent*. He simply argues that **Necessitism** should be preferred over **Contingentism** on theoretical grounds.

[12] A more traditional way to phrase the issue is in terms of *actualism* and *possibilism*. Roughly actualism is the view that only actually existing things exist, while possibilism is the view that some merely possibly existing things exist [96]. For the relation between the contingentism-necessitism debate and the actualism-possibilism debate, see [136, p. 22–25].

the property of being a man.[13] Thus, it follows from $\Diamond\neg$(Socrates exists) that $\Diamond\neg$(Socrates is a man). But by the interdefinability of possibility and necessity this is equivalent to $\neg\Box$(Socrates is a man), and so we would get by **SMA** and *modus tollens* that being a man is not an essential property of Socrates—contrary to our previous assumption that being a man is indeed a *paradigmatic* example of an essential property of Socrates. In other words, **SMA** is inconsistent with our background assumptions, while **MA**, as we've said before, agrees with our background assumptions. For this reason we will use **MA** rather than **SMA** in the following.

But back to Quine. Quine was, of course, no essentialist. In particular, he was skeptical of the idea of necessity pertaining directly to objects and their properties. Note that we have used the modal operator $\Box$ as an operator on *open formulas*: for example, **MA** contains the open formula $\Box$(if x exists, then x exemplifies Φ). Moreover, in **MA**, we *quantify into* this open formula: **MA** says that *for all properties* Φ *and all objects* x, Φ is an essential property of x iff $\Box$(if x exists, then x exemplifies Φ). Thus, the ranges of the quantifiers "for all properties Φ" and "for all objects x" extend into the open formula "$\Box$(if x exists, then x exemplifies Φ)," which is governed by the modal operator $\Box$. In other words, in **MA**, we quantify into a context of necessity. This gives us a precise sense in which the principle **MA** involves necessity pertaining to objects and their properties. In philosophical jargon, this notion of necessity, where it is possible to quantify into contexts of necessity, is called necessity *de re* ('of things'). Indeed, it seems that for a workable modal analysis of essential properties, we really *need to* quantify into the context of necessity: how else should we formalize the intuitive notion that essential properties are simply properties that are weakly necessary to their bearers?

In contrast, Quine [116] argued that the only way to make sense of necessity is as a concept that pertains to *sentences*. On this conception, necessity is simply the property of being true under all interpretations or *necessary truth*. In philosophical jargon, this notion of necessity is known as necessity *de dicto* ('of words').[14] Correspondingly, Quine argues that instead of the modal operator $\Box$, we should use the *sentential predicate* **nec** to formalize necessity. For a formula φ, Quine writes $\mathbf{nec}(\ulcorner``\varphi"\urcorner)$ with the intended in-

[13]Here we assume that an object can only exemplify a property if it exists. In more formal terminology, we assume that our background logic is a *negative free logic* [79].

[14]In modern contexts, necessity *de dicto* is also sometimes defined as a property of the contents of sentences or *propositions* [110, p. 9–13]. Quine himself was also skeptical of propositions [cf. 117], and here we mainly take Quine as our opponent. Thus, we'll only talk of necessity *de dicto* as a property of sentences. However, everything we say can easily be translated to apply to necessity *de dicto* as a property of propositions. For a more comprehensive discussion of the distinction between necessity *de dicto* and necessity *de re*, which also puts the distinction in a historical context, see [71].

terpretation that φ is necessarily true.[15] Note that for any formula φ, the formula $\mathbf{nec}(\ulcorner``\varphi"\urcorner)$ is a closed formula—a *sentence*—since the term $\ulcorner``\varphi"\urcorner$ is a closed term: it is a name of the formula φ. Consequently, in a formula $\mathbf{nec}(\ulcorner``\varphi"\urcorner)$, we cannot quantify over any variables that might occur free in φ—we cannot quantify into contexts of necessity [116, p. 170–71]. In particular, on Quine's approach, we may be able to write:

for all properties Φ and all objects x, Φ is an essential property of x iff **nec**("if x exists, then x exemplifies Φ"),

but this does not say the same thing as **MA**. Since the variables x and Φ in "**nec**("if x exists, then x exemplifies Φ")" are only *mentioned* and not *used*, they are not "captured" by the quantifiers "for all properties Φ" and "for all objects x." As a consequence, "**nec**("if x exists, then x exemplifies Φ")" simply says that the open formula "if x exists, then x exemplifies Φ" is necessarily true. But on the standard reading of open formulas, "if x exists, then x exemplifies Φ" is true under an interpretation iff *all objects* that satisfy "x exists" under the interpretation, together with *all properties* satisfy the formula "x exemplifies Φ" under the interpretation. In other words, "**nec**("if x exists, then x exemplifies Φ")" says that under all interpretations all existing objects exemplify all properties. Thus, the "analysis" of essential properties in terms of **nec** says that for all properties Φ and all objects x, Φ is an essential property of x iff under all interpretations all existing objects exemplify all properties—which is not only absurd, but certainly not what **MA** says. Indeed, it seems that on Quine's predicate approach to necessity *de dicto* there is no way to properly formulate a modal analysis of essential properties, since on the approach we can't quantify into contexts of necessity.[16] And so, since Quine held that the only way to make sense of necessity is as necessity *de dicto* which should be formalized using **nec**, Quine argued that we can't give a proper analysis of essential properties—he argued that the concept is confused.

Let's sum up Quine's criticism. On the one hand, Quine held that we *do* have a firm understanding of necessity *de dicto*: it amounts to the notion of a sentence being true under all interpretations.[17] And, so Quine, this notion should be formalized using the sentential predicate **nec**. But on this

[15] For more on the *Quine-corners* $\ulcorner\ \urcorner$, see [115, p. 36].

[16] But compare the recent developments in predicational treatments of modalities [59, 81, 57].

[17] A possible refinement may be to say that a sentence is necessary *de dicto* iff it is true under all interpretations *from a certain class of interpretations*. By restricting ourselves in this way to interpretations from a given class of interpretations, we can model different concepts of necessity *de dicto*: for example, if we're able to define the concept of a metaphysically admissible interpretation, for example, we can define metaphysical necessity *de dicto* as truth under all metaphysically admissible interpretations; or we could define logical necessity *de dicto* as truth under *all* interpretations.

predicate approach to necessity, it seems impossible to give a formulation of **MA** that adequately captures the intuition that essential properties are simply properties that are weakly necessary to their bearers, since on the approach we can't quantify into contexts of modality. On the other hand, so Quine, we *don't* have a clear notion of necessity *de re*. Moreover, to properly formulate the principle **MA**, we need to quantify into contexts of necessity. This is possible using the modal operator $\Box$, but, so Quine, the approach would commit us to a concept of necessity *de re*. Since Quine argued that we don't have a clear understanding of necessity *de re*, he concluded that also the concept of essential properties, and thus essentialism, is hopelessly muddled and confused.[18] And mainstream analytic metaphysics followed him in this assessment. At least for a while.

Things changed in the second half of the twentieth century with the rise of modal metaphysics. The rise of modal metaphysics "piggy-backed," as it were, on breakthrough results in the semantics of necessity. In groundbreaking work, Kripke [74, 78] developed an intuitively plausible semantics for necessity *de re* in terms of *possible worlds*.[19] There are different ways of developing a Kripke-style semantics, but here we shall informally sketch the semantics of [78], since this semantics is particularly suited to our contingentist essentialist needs.

Kripke [78] defines a class of structures that can interpret statements of necessity *de re*. Here we shall only informally sketch this kind of structure, leaving out many details that are not important to the present purpose. The central concept of Kripke's semantics is the concept of possible worlds. Intuitively, we can think of possible worlds as entities that correspond to the ways the world could have been.[20] In metaphysics, we standardly assume that for every way the world could have been, there is a possible world that corresponds to this way.[21] For example, since we've assumed that Socrates

[18]It is not entirely clear in [117] if Quine rejects essentialism because he rejects necessity *de re* or if he rejects necessity *de re* because he rejects essentialism. But for the present point this doesn't matter much: in any case, Quine rejected both because he thought we don't have a clear enough understanding of the concepts involved. For a more comprehensive discussion of Quine's views about modality, see [89] or [52].

[19]There is a historical debate on Kripke's priority with regard to the semantics. Contenders for priority include Bayart [6, 7], Carnap [19, 18], Hintikka [60], Kanger [67], Montague [100], and Prior [114] and others. As evidenced by the length of this list, the idea of using possible worlds for the semantics of modality was very much "in the air" at the time. Here we stick with the standard view, which gives Kripke priority. For an overview of the history of modal logic and its semantics, see [55] or [5]. For more comprehensive discussions of the relationship between quantified modal logic and modality *de re*, see, e.g., [91, 109, 93, 4].

[20]Here we wish to remain neutral with regard to the ontological nature and status of these entities. For a comprehensive discussion of the concept of possible worlds and the different views about their ontological status, see [97].

[21]This is known as *the principle of plenitude* [83, p. 86–92]. On an ontologically innocu-

and Xanthippe could have never met and thus could have never married, there is a possible world where Socrates is not married to Xanthippe. And since we've assumed that Socrates could have failed to exist and thus not have been a man, there is a world where Socrates doesn't exist and thus isn't a man. In contrast, since we've assumed that Socrates could not have existed without being a man, there is no possible world where Socrates exists but is no man. The actual world, the world that we live in, is the world that corresponds to the way the world is actually like. At the actual world, of course, Socrates and Xanthippe both exist, they are indeed married, and Socrates is a man.

Thus, intuitively, there are possible worlds where Socrates exists and worlds where he doesn't. More generally, for every world there will be a set of things that exist at that possible world—the *domain* of the world—and these domains can change from world to world. For this reason, the kind of semantics that we're talking about here is typically called a *variable domain* semantics.[22] Let's denote the set of all possible worlds by $\mathcal{W}$. Then, for every world $w \in \mathcal{W}$ there will be a set $\mathcal{D}_w$ of things that exist at that possible world: for all objects x and all worlds $w \in \mathcal{W}$, x exists at w iff $x \in \mathcal{D}_w$.

Some aspects of Kripke's semantics will be important in the following: First, we assume that the quantifiers "for all objects" and "for some objects" have existential import: both expressions range only over the *existing* objects. Thus, it's sufficient for it to be the case at a possible world $w \in \mathcal{W}$ that all objects are men that all the objects in $\mathcal{D}_w$ exemplify the property of being a man and it's sufficient for some objects at $w \in \mathcal{W}$ to be married to Xanthippe that some object in $\mathcal{D}_w$ is married to Xanthippe.[23] Second, individual constants, like "Socrates," get assigned a fixed denotation within $\bigcup_{w \in \mathcal{W}} \mathcal{D}_w$. For example, we'll assume that the denotation of "Socrates" is Socrates, the man. In philosophical jargon, we treat individual constants as *rigid designators* [76]. Intuitively, this requirement makes sure that when we talk about Socrates in a sentence like "Socrates is a man," we are really talking about Socrates and not some arbitrary object that is named "Socrates." Third, in contrast to individual constants, our interpretation of predicates, like "x is a man" or "x exists," can change from world to world. Intuitively, the things that have a property at a world can change from world to world. Thus, the set of things that have a property at a world—the *extension* of the property at the world—can change. Formally, we assign to every predicate, like "x is a man" or "x exists," an extension at every possible world. For example, the extension that we assign to "x is a man" at a world $w \in \mathcal{W}$ is the set

ous reading, the principle is relatively uncontroversial [cf. 97, § 2.1.4].

[22] For a more precise development and discussion of these issues, see [51, p. 101–5, 163–185] and [62, p. 252–72].

[23] This ties in with our assumption that our background logic is a negative free logic, see footnote 13.

$\{x \mid x \text{ is a man at } w\}$. And the extension that we assign to "x exists" at a world $w \in \mathcal{W}$ is the set $\{x \mid x \text{ is exists at } w\} = \mathcal{D}_w$.[24]

Now, Kripke's insight was that we can understand what is necessary *de re* for an object as what is the case for the object in all possible worlds [74, 78].[25] Based on this idea, Kripke developed a semantics for statements of necessity *de re*.[26] On this semantics, a statement of necessity *de re* is true iff in all possible worlds it is the case what the statement says is necessary *de re*:

Kripke-$\Box$. For all formulas φ, the statement $\Box\varphi$ is true iff what φ says is the case at every possible world.

So, for example, the statement "$\Box$(Socrates is married to Xanthippe)" is true on Kripke's semantics iff at every possible world Socrates is married to Xanthippe. Indeed, given what we just said about possible worlds, the statement is false. In contrast, the statement "$\Box$(if Socrates exists, then Socrates is a man)" is true on Kripke's semantics iff in every possible world where Socrates exists, he is a man. Indeed, given what we just said about possible worlds, the statement is true.

The semantics can, of course, also be extended to statements of possibility *de re*—possibility that applies directly to things and their properties. We simply say that a statement of possibility *de re* is true iff there is some possible world where what the statement says is possible *de re* is the case.

Kripke-$\Diamond$. For all formulas φ, the statement $\Diamond\varphi$ is true iff what φ says is the case at some possible world.

So, for example, the statement "$\Diamond$(Socrates exists & $\neg$(Socrates is a man))" is true under Kripke's semantics iff there is some possible world where Socrates exists but does not exemplify the property of being a man. Indeed, given our assumptions about possible worlds, the statement is false. In contrast, the statement "$\Diamond$(Socrates exists & $\neg$(Socrates is married to Xanthippe))" is true on Kripke's semantics iff there is some possible world where Socrates exists but is not married to Xanthippe. Indeed given our previous assumptions about possible worlds, the statement is true.

Note that Kripke's semantics validates our standard assumptions about

---

[24] Note that a consequence of this semantics is that an object can only be in the extension of a predicate at a world if the object exists at that world, see footnote 13.

[25] In fact, it is also possible to define necessity *de dicto* in this framework: we simply say that a sentence is necessary *de dicto* iff it is true *at* all possible worlds. Since possible worlds intuitively correspond to ordinary interpretations, this is roughly the same idea as what we sketched in footnote 17. Here, however, we're mainly interested in necessity *de re* and will not discuss necessity *de dicto* in more detail.

[26] As we've pointed out already, here we will only give an informal account of the semantics. For a more detailed and formal development, see [51, 62].

modal logic. Remember that we've assumed that everything that is necessarily the case is actually the case. Indeed on Kripke's semantics the following holds for all statements φ:

$$\text{If } \Box\varphi \text{ is true, then } \varphi \text{ is true.}$$

To see this, note that, by **$\Box$-Kripke**, $\Box\varphi$ is true iff what φ says is the case at every possible world. But we've said that the actual world is also a possible world and thus at the actual world it's the case what φ says. In other words, φ is true.[27] Also, by the duality of the quantifiers "for some" and "for all" and some standard assumptions about negation, the semantics validates the interdefinability of possibility and necessity: for all statements φ,

$$\Diamond\varphi \text{ is true iff } \neg\Box\neg\varphi \text{ is true.}$$

To see this note that, by **$\Diamond$-Kripke**, $\Diamond\varphi$ is true iff what φ says is the case at some possible world. Now, at this possible world it certainly isn't the case what $\neg\varphi$ says, since, intuitively, $\neg\varphi$ says that what φ says is not the case. Thus, it's not the case that at every possible world what $\neg\varphi$ says is the case. Hence, by **$\Box$-Kripke**, $\Diamond\varphi$ is true iff $\Box\neg\varphi$ is *not* true. But then, since we may standardly assume for all statements ψ that $\neg\psi$ is true iff ψ is not true, we get, that $\Diamond\varphi$ is true iff $\neg\Box\neg\varphi$ is true. Indeed, Kripke [78] showed that, in a formally precise sense, the semantics we just sketched is *sound and complete* for the standard modal logic $S5$, which includes, among others, our principles **T** and $\Diamond/\Box$.[28]

Moreover, note that Kripke's semantics satisfies **Contingentism**, in the sense that on Kripke's semantics, given our assumptions about possible worlds, we get that:

$$\text{“}\Diamond(\text{for some } x\ \Diamond\neg(x \text{ exists}))\text{” is true.}$$

To see this, first note that we've assumed that there is a possible world where Socrates doesn't exist. Thus, by **$\Diamond$-Kripke**, the statement "$\Diamond\neg$(Socrates exists)" is true. But we've assumed that, at the actual world, Socrates exists. Thus, at the actual world there is some object x, namely Socrates, such that $\Diamond\neg(x$ exists). But since the actual world is also a possible world, there is a world, namely the actual world, where it's the case what "for some object x $\Diamond\neg(x$ exists)" says. Hence, by **$\Diamond$-Kripke** again, "$\Diamond$(for some x $\Diamond\neg(x$ exists))"

[27] Here we assume, of course, that a statement is true iff what it says is actually the case. This assumption, however, is a fairly standard informal gloss of the fairly uncontroversial *T-scheme*. Reasoning in the context of the semantic paradoxes might lead us to abandon the view that this principle holds for *all* statements, but, in any case, the statements we're talking about here are fairly safe.

[28] For an argument that $S5$ is indeed the best way to treat metaphysical necessity, see [136, p. 92–119, 130–39]. For a precise proof of the theorem that $S5$ is sound and complete with respect to Kripke's semantics, see [51, 62].

is true.[29]

The success of Kripke's semantics single-handedly dispersed Quine's worries with essentialism. Remember that Quine's argument against essentialism mainly rests on the assumption that we don't have a clear concept of necessity *de re*. But Kripke's semantics *does* give us a precise and intuitively plausible understanding of necessity *de re* in terms of possible worlds. Moreover, Kripke's semantics allows us to embed the concept of essential properties into the framework of possible worlds: By translating **MA** via **Kripke-**$\Box$ we get an analysis of essential properties in terms of possible worlds:

Possible Worlds Analysis (PWA). For all properties Φ and objects x,

> Φ is an essential property of x iff at every possible world where x exists, x exemplifies Φ.

Thus, using Kripke's semantics, the framework of possible worlds allows us to give a precise analysis of essential properties. The framework of possible worlds is indeed quite powerful in that it allows us to analyze a wide range of philosophical concepts [110, 76, 83]. In short, possible worlds are the perfect playground for metaphysicians. And indeed, Kripke's work sparked the rise of *modal metaphysics*: the paradigm of approaching metaphysical questions mainly in terms of modality and possible worlds. For most of the second half of the twentieth century, modal metaphysics was the predominant paradigm in analytic metaphysics and **MA** under its reading in Kripke's semantics—**PWA**—was the predominant analysis of the concept of essential properties [109, 90, 110, 76, 83].

Part of the reason why **PWA** became the standard analysis of essential properties is that the possible worlds framework is, metaphysically speaking, a great "package deal." In particular, the possible worlds framework does not only allow us to analyze the concept of *essential* properties, but we can also use it to analyze the notion of a *property* simpliciter. Given what we said above, we essentially interpret predicates by means of what Carnap [19] calls *intensions*: functions that assign extensions to possible worlds. Since we express properties by means of predicates—for example, we express the property of being a man by means of the predicate "x is a man"—this semantics suggests *identifying* properties with intensions [99, 83]:

Intensional Property Theory (IPT). For all Φ, Φ is a property iff Φ is a function that assigns to every possible world $w \in \mathcal{W}$ the set $\Phi(w) \subseteq \mathcal{D}_w$ of things that exemplify the property at that world.

On this theory, then, we can then analyze the relation of an object exemplifying a property at a world as the object being an element of the extension

[29] For a discussion of whether the semantics is intuitively adequate for *actualism*, the view that only actual things exist, see [96, §3.3]

of the property at the world:

Exemplification in IPT. For all objects x, for all properties Φ, and for all worlds $w \in \mathcal{W}$, x exemplifies Φ at w iff $x \in \Phi(w)$.

To illustrate, on **IPT**, we would take the property of being a man, for example, to be the function $\text{being\_a\_man} : \mathcal{W} \to \wp(\bigcup_{w \in \mathcal{W}} \mathcal{D}_w)$ such that $\text{being\_a\_man}(w) = \{x \in \mathcal{D}_w \mid x \text{ is a man at } w\}$. Thus, by **Exemplification in IPT**, we get that Socrates exemplifies the property of being a man at a world $w \in \mathcal{W}$ iff Socrates $\in \text{being\_a\_man}(w)$. And since $\text{being\_a\_man}(w) = \{x \in \mathcal{D}_w \mid x \text{ is a man at } w\}$, we get that Socrates exemplifies being a man at a world iff Socrates is a man at the world.

Let's sum up the results of this section. We've discussed the most popular analysis of the concept of essential properties throughout twentieth century analytic philosophy: the *modal analysis* **MA**, which analyzes essential properties in terms of weak necessity *de re*. We've seen how this analysis was historically vindicated against Quine's attacks. Quine objected that the notion of necessity *de re* is confused, but Kripke gave us a clear understanding of necessity *de re* in terms of possible worlds: **MA** was vindicated by translating it via Kripke's semantics into the possible worlds analysis **PWA**. Moreover, in the possible worlds framework **PWA** was supplemented with the *intensional property theory* **IPT** to give us a full-blown analysis of the concept of essential properties in the possible worlds framework: Together, **PWA** and **IPT** give us a metaphysically robust philosophical theory of essential properties, where all concepts involved are clearly explicated. Moreover, all of this can be carried out in such a way that we satisfy **Contingentism**: the view that it's possible for some things to possibly not exist or, in short, that ontology is contingent. This was the state of the art—at least until the 90s.

1.2 Counterexamples to the Modal Analysis

Toward the end of the twentieth century, the modal analysis of essential properties was beginning to be called into question. Fine [39] famously came up with a range of counterexamples to the modal analysis. He discusses the modal analysis quite generally in various forms, but here we shall focus on Fine's criticism as it applies to the possible worlds analysis **PWA** supplied with the intensional property theory **IPT**.

Fine's counterexamples can be grouped into four categories according to the concepts that they involve: (1) set membership, (2) identity and distinctness, (3) necessary truths, and (4) existence [22, p. 64]. All of the counterexamples have in common that they purport to show that **PWA** does not always agree with our intuitions about essential properties. More specifically, in each of the examples, the analysans—weak necessity *de re*—is present, while intuitively the analysandum—the concept of essential properties—is not. In other words, the counterexamples purport to show it's not *sufficient* for a property Φ to be an essential property of an object x that x exemplifies Φ in every possible world where x exists.[30] Let's go through these examples in turn:[31]

(1) **Set Membership.** Philosophers often assume that for every set it's necessary *de re* that the set exists iff all of its members do [40, 108]. Indeed, given that sets are typically said to be individuated by their members and that we may for all pluralities of objects form the set of these objects, from the perspective of a modal metaphysician, this assumption is quite plausible. It's somewhat tedious to spell out this principle in full generality, since sets may have arbitrarily infinitely many members.[32] So, instead, let's illustrate the consequence of the principle in the case of Fine's example of Socrates and the set {Socrates}—Socrates' singleton—whose sole member is Socrates. By the principle that for every set it's necessary that the set exists if and *only if* all of its members do, we get for Socrates and his singleton that:

    - for every possible world $w \in \mathcal{W}$ such that $\{\text{Socrates}\} \in \mathcal{D}_w$, we have that Socrates $\in \mathcal{D}_w$ and Socrates is a member of {Socrates} at w.

    Now, we may consider the property of being such that Socrates exists and having him as a member. By **IPT**, this property exists, indeed we may simply identify it with the function which maps a world $w \in \mathcal{W}$ to an extension in $\mathcal{D}_w$ according to the rule:

    $$w \mapsto \{x \in \mathcal{D}_w \mid \text{Socrates} \in \mathcal{D}_w \;\&\; \text{Socrates is a member of } x \text{ at } w\}.$$

    Then, by **PWA**, we immediately get that this property is an essential property of {Socrates}—as it should be.

[30] Fine [39] argues, however that the *converse* direction holds: if a property Φ is essential to an object x, then at every possible worlds where x exists, x exemplifies Φ. In other words, weak necessity *de re* is *necessary* for essential properties.

[31] Dunn [34] already gave some of the examples that Fine gave in his [39]. However, Dunn's point is slightly different from Fine's. Dunn aims to show that some relational predicates only determine essential properties in some of their positions. However, the conclusions he draws from these examples are very similar to Fine's. Since the issue of priority is only marginally important to the problem at hand—the (purported) failure of the modal analysis—we shall discuss [34] and his point only in the footnotes of this section.

[32] If we were to allow arbitrarily infinitary quantifiers, we could formulate the principle by saying that for all objects $x_1, x_2, \ldots$ and all sets y, $\Box$(if $x_1, x_2, \ldots$ are all and only the elements of y, then $\Box$(y exists iff all of $x_1, x_2, \ldots$ exist). However, arbitrarily infinitary quantification in the sense that is required for this principle is technically speaking difficult to handle, to say the least [10].

However, given **PWA** and **IPT**, the principle about sets also has counterintuitive consequences. Note that if we say that for any set it's necessary the set exists *if* and only if all of it's members do, we get for Socrates and his singleton that:

- for every possible world $w \in \mathcal{W}$ such that Socrates $\in \mathcal{D}_w$, we have that {Socrates} $\in \mathcal{D}_w$ and Socrates is a member of {Socrates} at w.

Now consider the property of being such that {Socrates} exists and being a member of {Socrates}. Again, on **IPT**, this property exists, as we may simply identify it with the function which maps a world $w \in \mathcal{W}$ to an extension in $\mathcal{D}_w$ according to the rule:

$w \mapsto \{x \in \mathcal{D}_w \mid \{\text{Socrates}\} \in \mathcal{D}_w \;\&\; x \text{ is a member of } \{\text{Socrates}\} \text{ at } w\}$.

And **PWA**, it immediately follows that this property is an essential property of Socrates.

But, intuitively, this is not plausible. As Fine puts it: "There is nothing in the nature of a person, if I may put it this way, which demands that he belongs to this or that set or which demands, given that the person exists, that there even be any sets" [39, p. 5]. Thus, Fine argues that essential properties should intuitively capture the *nature* of their bearers [39, p. 1], and since there is intuitively nothing in the nature of Socrates that connects him to his singleton, or any other set for that matter: it's neither an essential property of Socrates that his singleton exists nor that Socrates is a member of it.[33]

(2) **Identity and Distinctness.** Philosophers often assume that for any two identical objects it's necessary that the two objects are identical whenever they both exist and that for any two distinct objects it's necessary that they are distinct whenever they both exist. These two principles are known as the *necessity of identity* and *the necessity of distinctness* respectively [75]:

Necessity of Identity. For all objects x and y, if x is identical to y, then for all possible worlds $w \in \mathcal{W}$, if $x, y \in \mathcal{D}_w$, then x is identical to y at w.

[33] Dunn [34, p. 90–91] gives a very similar example, just that he talks about the pair set {Tom, Dick}. As Fine, Dunn argues that it's an essential property of {Tom, Dick} that it contains Tom and Dick as members, but that it's *not* an essential property of either Tom or Dick to be members of {Tom, Dick}. The phenomenon that Dunn wishes to illustrate with this example is that some predicates, such as the membership predicate "∈," only express essential properties in some of their positions: the membership predicate, for example, generally expresses an essential property in its second position, but not always in it's first. Dunn calls situations like this cases of *asymmetric* essence. Note that Fine [39, p. 5] uses a very similar terminology. Given all of this, it seems that priority for this counterexample belongs to Dunn.

Necessity of Distinctness. For all objects x and y, if x is distinct from y, then for all possible worlds $w \in \mathcal{W}$, if $x, y \in \mathcal{D}_w$, then x is distinct from y at w.

Indeed, if we understand identity as *numerical* identity and distinctness as *numerical* distinctness, these principles are nearly trivially true on the version of the possible worlds framework we've sketched above.[34]

The principles **Necessity of Identity** and **Necessity of Distinctness** capture the intuition that objects are essentially what they are and that they are essentially not what they are not. To illustrate, let's consider Fine's example of Socrates and the Eiffel Tower. Since it's of course actually the case that Socrates is identical to Socrates, we get by **Necessity of Identity** that:

- for all possible worlds $w \in \mathcal{D}_w$ such that Socrates $\in \mathcal{D}_w$, Socrates is identical to Socrates at w.

Now, let's consider the property of being self identical. According to **IPT**, this property exists, as we may simply identify it with the function which maps a world $w \in \mathcal{W}$ to an extension in $\mathcal{D}_w$ according to the rule:

$$w \mapsto \{x \in \mathcal{D}_w \mid x \text{ is identical to } x \text{ at } w\}.$$

Then, by **PWA**, this property is an essential property of Socrates. In other words, it's an essential property of Socrates to be who he is—as it should be intuitively.

In contrast, since it's of course actually the case that Socrates is distinct from the Eiffel Tower, we get by **Necessity of Distinctness** that:

- for all possible worlds $w \in \mathcal{D}_w$ such that Socrates, the Eiffel Tower $\in \mathcal{D}_w$, Socrates is distinct from the Eiffel Tower at w.

Now consider the property of being such that the Eiffel Tower exists and being distinct from the Eiffel Tower. According to **IPT**, this property exists, as we may simply identify it with the function which maps a world $w \in \mathcal{W}$ to an extension in $\mathcal{D}_w$ according to the rule:

$$w \mapsto \{x \in \mathcal{D}_w \mid \text{Eiffel Tower} \in \mathcal{D}_w \ \&\ x \text{ is distinct from the Eiffel Tower at } w\}.$$

Then, by **PWA**, this is again an essential property of Socrates. But this, so Fine, is not intuitively plausible: "But it is not essential to Socrates that he be distinct from the Tower; for there is nothing in his nature

[34] Moreover, the syntactic formulations of **Necessity of Identity** and **Necessity of Distinctness** in terms of $\Box$ are *logically equivalent* in the modal logic that the framework determines. However, we shall keep the two principles apart because we wish to focus on their individual *metaphysical* consequences, which, it turns out, are quite different.

which connects him in any special way to it" [39, p. 5]. Thus, as in the case of Socrates and his singleton, the counterexample rests on the intuition that Socrates's nature is not connected to the Eiffel Tower, and hence, intuitively, being distinct from the Tower if it exists is not among Socrates essential properties. Indeed, the same argument can be given for *any* object distinct from Socrates.

(3) **Necessary Truths.** Most philosophers assume that there are necessary truths. For example, many philosophers assume that the truths of mathematics are necessary [76, 107]:

Necessity of Mathematics. For all mathematical statements φ, if φ is true, then $\Box\varphi$ is true.

To illustrate, consider the statement "there are infinitely many prime numbers". By Euclid's theorem, we know that this statement is true. And since it's undoubtedly a statement of mathematics, we get by **Necessity of Mathematics** that "$\Box$(there are infinitely many prime numbers)" is true. And on Kripke's semantics, by the clause **$\Box$-Kripke**, this is the case iff at every possible world there are infinitely many prime numbers. More generally, given **Necessity of Mathematics**, it will be the case at every possible world what any true statement of mathematics says.

Moreover, many philosophers assume that for every statement φ there is a property of being such that what φ is the case. Since what a statement says—its *content*—is also called a *proposition*, such properties are typically called *propositional properties* [130, §7.5]:

Propositional Properties. For every statement φ, there is a property of being such that what φ says is the case. And for all objects x and all statements φ, x exemplifies being such that what φ says is the case iff x exists & what φ says is the case.

Indeed, on the intensional property theory **IPT**, we can show that there are propositional properties: For a statement φ, simply take the property of being such that what φ says is the case to be the function which maps a world $w \in \mathcal{W}$ to an extension in $\mathcal{D}_w$ according to the rule:

$$w \mapsto \begin{cases} \mathcal{D}_w & \text{if what } \varphi \text{ says is the case at } w \\ \emptyset & \text{otherwise} \end{cases}$$

And by **Exemplification in IPT**, we get that an object exemplifies the property of being such that what φ says is the case at a world iff what φ says is the case at the world.

So, for example, by **Propositional Properties**, there is a property of being such that what "there are infinitely many prime numbers" says is the case—the property of being such that there are infinitely many prime numbers. And an object exemplifies this property iff what "there are infinitely many prime numbers" says is indeed the case—iff there are infinitely many prime numbers. But, as we've just discussed, by **Necessity of Mathematics**, it's necessarily the case that there are infinitely many prime numbers. And that means that it's necessary *de re* for *any* object that exists to exemplify the property of being such that there are infinitely many prime numbers. In particular, we get:

- for all possible worlds $w \in \mathcal{W}$, if Socrates $\in \mathcal{D}_w$, then Socrates exemplifies being such that there are infinitely many prime numbers at w.

And, by **PWA**, this means that the propositional property of being such that there are infinitely many prime numbers is an essential property of Socrates. Indeed, for *any* necessarily true proposition, according to **PWA**, the corresponding propositional property is going to be an essential property of *any* object whatsoever. This is certainly counterintuitive. As Fine puts it: "[I]t is no part of Socrates' essence that there be infinitely many prime number" [39, p. 5]. Since Fine uses "essence" and "nature" interchangeably, the issue is again that Socrates' nature is not connected to the prime numbers and thus, intuitively, there being infinitely many prime numbers is not among his essential properties.

(4) **Existence.** Finally, some authors assume that *existence* is itself a property [13]. Indeed, on the intensional property theory **IPT**, we can simply identify this property with the function which maps a world $w \in \mathcal{W}$ to an extension in $\mathcal{D}_w$ according to the rule:

$$w \mapsto \mathcal{D}_w.$$

By **Exemplification in IPT**, it follows immediately that an object exemplifies this property at a world iff the object is a member of the domain at the world—in other words, an object exemplifies the property at world iff the object exists there.[35] Now notice that plugging existence into **PWA**, we get that:

- for all objects x, existence is an essential property of x iff for all possible worlds $w \in \mathcal{W}$ such that $x \in \mathcal{D}_w$, we have $x \in \mathcal{D}_w$.

But this is trivially the case! Thus, by **PWA**, existence is an essential property of *any* object whatsoever. More specifically, we get that:

- for all possible worlds $w \in \mathcal{W}$ with Socrates $\in \mathcal{D}_w$, we have that Socrates $\in \mathcal{D}_w$.

[35] For a critical discussion of the assumption that existence is a property, see [102].

> Thus, by **PWA**, we get that existence is an essential property of Socrates. But this is, so Fine, intuitively implausible. As he simply puts it: "we do not want to say that he essentially exists" [39, p. 6].

Note that all of Fine's counterexamples involve the informal claim that in order for a property Φ to be an essential property of an object x the property Φ has to be *part of the nature* of the object x. Let's grant Fine this assumption for now. But why should we believe that facts about {Socrates} (1), about the Eiffel Tower (2), about the prime numbers (3), and about Socrates' existence (4) are *not* part of the nature of Socrates? In the cases (1–3), in order to support the claim that the facts are not part of Socrates' nature, Fine gives an *epistemological* argument: For assume that facts about {Socrates}, about the Eiffel Tower, or about the prime numbers were part of the nature of Socrates. Then it would not be possible to discover Socrates' nature without learning about all of these objects. But surely it is possible to understand the nature of Socrates without learning about sets, the Eiffel Tower, or prime numbers. Otherwise, as metaphysicians, we would be in the quite uncomfortable (or comfortable, depending on the perspective) situation of having to say something about all of these things in order to fully describe the nature of Socrates. As Fine quips: "O happy metaphysician! For in discovering the nature of one thing, he thereby discovers the nature of all things" [39, p. 6].[36] Now, Fine might be stressing his point a little bit too much when he says that we have to learn about the *nature* of all of these things, but given what we've just said, it's certainly the case that we have to say *something* about them: namely, that {Socrates} contains Socrates, that the Eiffel Tower is distinct from Socrates, and that there are infinitely many prime numbers. To make things worse, as we've noted, by analogous arguments to the ones given above, we can extend this to *all* sets that contain Socrates, to *all* objects distinct from Socrates, and to *all* necessary truths. Certainly, it's not feasible to talk about all these things, if we simply wish to describe Socrates' nature. In short, Fine gives an epistemic *reductio* argument for his claims on the possibility of a feasible metaphysics *assuming* that the essential properties of an object have to be part of its nature.

In the case of Socrates' existence (4), the epistemological argument does not work so well: Socrates' existence does not involve any objects distinct from Socrates, and thus, even if we took existence to be part of Socrates' nature, we could describe Socrates nature without having to talk about an unfeasible number of objects distinct from Socrates. Admittedly, it might seem *odd* to say that it's part of the nature of all objects that they exist, but this judgement appears to rely on a specific understanding of what constitutes

[36]Hannes Leitgeb remarks that this in fact sounds like something Leibniz would have proposed. However, here we shall not pursue this (mostly historical) claim and rest with the assumption that as non-Leibnizians this is indeed implausible.

the nature of an object. But a proponent of **PWA** is not committed to such an understanding. For example, a modal metaphysician in favor of **PWA** could say that the nature of an object contains all and only those properties that are directly relevant to the existence of the object. Then, existence would merely be a limit case: obviously, for any object, existence is trivially directly relevant to the existence of the object. Indeed, the modal fact in virtue of which **PWA** classifies existence as an essential property of any object whatsoever is a *logical* necessity: the fact that it's necessary *de re* for ever object to exist, if it exists. This really just seems to be a limit case of being part of an object's nature in the above sense. Thus, Fine's claim that "we do not want to say that [Socrates] essentially exists" [39, p. 6] is not very convincing from a modal metaphysician's perspective. If, however, we think of the nature of an object as something that *defines* the object or the like [39, p. 10–15], then it's indeed plausible that existence is not part of the nature of an object: clearly, existence is not part of the definition of Socrates.

Note further that all of Fine's counterexamples involve more or less contentious assumptions. Fine is well-aware of this. He writes:

> *[It is not] critical to the example[s] that the reader actually endorse the particular modal and essentialist claims to which I have made appeal. All that is necessary is that he should recognize the intelligibility of a position which makes such claims. For any reasonable account of essence should not be biased towards one metaphysical view rather than the other. It should not settle, as a matter of definition, any issue which we are inclined to regard as a matter of substance.* [39, p. 5]

Thus, Fine's point is that the analysis of a metaphysical concept, like the concept of essential properties, should not commit us to a specific metaphysical view. But if we wish to subscribe to the modal analysis **PWA** and at the same time believe that it's necessary that sets exist iff all of their members do, that it's necessary that distinct objects are distinct, or that it's necessary that mathematical truths hold, we seem to be committed to counterintuitive consequences. Thus, we have to abandon either these views or the modal analysis. And, by Fine's criterion that an analysis should not commit us to any specific metaphysical views, the thing to be abandoned is the modal analysis.

But, especially in the present context, this point doesn't seem to carry as much force as it might initially appear to. Remember, for example, that the reason why we chose **MA** over **SMA** as our working modal analysis of essential properties is that we wanted to subscribe to **Contingentism** rather than **Necessitism**: on the analysis **SMA**, given some other assumptions, we we were not able to hold **Contingentism** and at the same time hold that being a man is an essential property of Socrates. Thus, it seems that **MA** is biased towards a specific metaphysical view, namely **Contingentism**. Moreover, we *wanted* it to be like this: our aim in this dissertation is to give an explication of the concept of essential properties *assuming* the thesis of **Contingentism**. To give another example, consider Carnap's famous argument against Heidegger's claim that *the Nothing itself nothings* [20]. Carnap argues that on what he thinks is the proper analysis of existence and non-existence by means of the existential quantifier, the claim is *meaningless* in one form or another. Thus, Carnap's analysis of existence and non-existence is (in some sense) biased against Heidegger's view: it renders it meaningless. Heidegger, of course, thought that the view that the Nothing nothings is deeply meaningful and indeed true. And a lot of (continental) philosophers agree with him on this [129]. The point here is that adopting an analysis of a concept, like **PWA**, in some conceptual framework, like the framework of possible worlds, nearly always renders some position false or even meaningless—most likely even a position that some opponent might hold to be meaningful and important. This is part of the point of adopting an analysis in some conceptual framework in the first place. Thus, the fact that the modal analysis **PWA** in the framework of possible worlds may be biased against certain metaphysical views is not *in itself* a problem.

It seems that if we grant Fine's assumption that it's a necessary condition for a property to be an essential property of an object that the property is part of the nature of the object, Fine's counterexamples leave us with two options: either we reject all the metaphysical assumptions that give rise to the counterexamples (1–3) or we abandon **PWA** together with **IPT** as our analysis of essential properties in the framework of possible worlds. But as modal metaphysicians working in the possible worlds framework, all the assumptions needed for the counterexamples are quite plausible: we can show that all of the properties involved exist by the intensional property theory **IPT** and our assumptions about possible worlds validate the problematic cases of necessity *de re*. Thus, the most obvious candidate for adjustment is the modal analysis **PWA**.

A natural way of trying to repair **PWA** in light of Fine's counterexamples is by imposing a *condition of relevance* on essential properties. Fine writes:

> *[O]ne might try to add a condition of relevance to the modal criterion. One would demand, if a property is to be essential to an object, that it somehow be relevant to the object.* [39, p. 6]

Remember that above we said that a modal essentialist could plausibly say that the nature of an object contains all and only the properties that are (in some sense) modally relevant to the existence of the object. Thus, if we assume that for a property to be an essential property of an object, the property has to be part of the nature of the object, we get the following condition:

Relevance Condition. For all properties Φ and all objects x, if Φ is an essential property of x, then Φ is wholly relevant to x.

Now, the notion of relevance in this condition will, of course, have to be made more precise, and we'll get back to that. But let's for the moment simply assume an informal notion of relevance to illustrate the point. It seems that indeed in all the example (1–3) we have a violation of the relevance condition: intuitively, the set fact that {Socrates} exists and Socrates is its member is not wholly relevant to Socrates, the fact that the Eiffel Tower is distinct from Socrates is not wholly relevant to Socrates, and there being infinitely many prime numbers is also clearly not relevant to Socrates. The reason why we might be inclined to say that these properties are not wholly relevant to Socrates is that they involve objects different from Socrates. Indeed, the failure of relevance in this sense seems to be exactly the reason why we're not intuitively inclined to classify the relevant properties as essential properties of Socrates. Thus, it seems that we have a philosophically motivated non-*ad hoc* answer to Fine's counterexamples.

Of course, as we've said, we'd have to make the informal notion of relevance in the **Relevance Criterion** philosophically precise for the answer to work. A straightforward way of doing so that immediately suggests itself is in terms of *relevance logic* (or relevan*t* logic, as more Anglo-influenced philosophers call it).[37] The idea would be that in **MA**, we take the conditional in the analysans $\Box$(if x exists, then x exemplifies Φ) to be a *relevant* conditional. Thus, we'd get something like the following of the modal analysis:

Relevant Modal Analysis (RMA). For all properties Φ and all objects x, Φ is an essential property of x iff $\Box$(if x exists, then, relevantly, x exemplifies Φ).

This analysis would then, once made more precise, almost trivially satisfy the **Relevance Condition**: if for some object x and property Φ, it's the case that $\Box$(if x exists, then, relevantly, x exemplifies Φ), then x exemplifying Φ has to be modally relevant to x existing. Of course, for this approach to work, we'll have to make the relevant conditional 'if ..., then, relevantly, ..." and the notion of a property being relevant to an object philosophically precise—for example, by means of a semantic analysis. But in any case, this seems like a promising way to go.[38]

Fine, however, argues that a relevance approach to the problem, as for example by **Relevance Condition** and **RMA**, is in serious trouble. He writes:

[37]For an overview of relevance logic, see [63, 92].

[38]Indeed, this is along the lines of what Dunn [34] suggests: He defines the notion of *necessary relevant exemplification* and suggests to analyze essential properties in just this way.

> *[T]he case of Socrates and his singleton makes it hard to see how the required notion of relevance could be understood without already presupposing the concept of essence in question. For we want to say that it is essential to the singleton to have Socrates as a member, but that it is not essential to Socrates to be a member of the singleton. But there is nothing in the "logic" of the situation to justify an asymmetric judgement of relevance; the difference lies entirely in the nature of the objects in question.* [39, p. 6–7]

The problem that Fine points out here is that in some cases we *do* wish to say that a property is an essential property of an object even though it involves an object distinct from it: Socrates is of course distinct from {Socrates}, but it's intuitively an essential property of the set to contain the man (if both exist). Indeed, the modal analysis captures this intuition in virtue the fact that it's necessary *de re* for all sets to exist iff all of their members do. But this very modal fact also falsely classifies the property of being contained in his singleton as an essential property of Socrates. Thus, the intuitive notion of a property being relevant to an object that we used above—the notion of a property not involving (in some sense) objects different from the one in question—does not work. Some properties are intuitively part of the nature of an object, and thus wholly relevant to the object, that involve objects distinct from the one in question: for example, the property of {Socrates} to contain Socrates as a member. And it is left unclear how this notion of a property being relevant to an object should be analyzed while avoiding these problems.

Intuitively, the problem here appears to be that the necessity *de re* that sets exist iff all of their members do somehow *arises* from the nature of the sets and not from the nature of the objects they contain. More generally, it seems a necessary condition for a property to be part of the nature of an object—or to be wholly relevant to the object—that having the property arises in this way from the nature of the object. But the notion of necessity *de re*, so Fine, is ill-suited to analyze this notion:

> *The concept of metaphysical necessity [...] is insensitive to source: all objects are treated equally as possible grounds of necessary truth; they are all grist to the necessitarian mill. What makes it so easy to overlook this point is the confusion of subject with source. One naturally supposes, given that a subject-predicate proposition is necessary, that the subject of the proposition is the source of the necessity. One naturally supposes, for example, that what makes it necessary that singleton 2 contains (or has the property of containing) the number 2 is something about the singleton. However, the concept of necessity is indif-*

ferent to which of the many objects in a proposition is taken to be its subject. The proposition that singleton 2 contains 2 is necessary whether or not the number or the set is taken to be the subject of the proposition. [39, p. 9]

In short, according to Fine, the concept of necessity *de re* is simply incapable of properly analyzing the notion of essential properties while satisfying the condition that an essential property of an object be part of the nature of the object.

We can push this point even further to get to the core of the problem. Take the two properties of being self-identical and being contained in one's singleton. It's easy to check that according to **IPT** both properties exist: the property of being self-identical is simply the function which maps a world $w \in \mathcal{W}$ to an extension in $\mathcal{D}_w$ according to the rule:

$$w \mapsto \{x \in \mathcal{D}_w \mid x \text{ is identical to } x\},$$

and the property of being contained in one's singleton is the function which maps a world $w \in \mathcal{W}$ to an extension in $\mathcal{D}_w$ according to the rule:

$$w \mapsto \{x \in \mathcal{D}_w \mid x \text{ is a member of } \{x\}\}.$$

By the principle that it's necessary *de re* that sets exist iff all their members do and **Exemplification in IPT**, it follows that any object exemplifies both properties in all worlds where it exists. More technically, two properties are said to be *necessarily equivalent* iff for all objects it's necessary *de re* that the object exemplifies the one property iff it exemplifies the other:

Necessary Equivalence (Definition). For all properties Φ and Ψ, the property Φ is necessarily equivalent to the property Ψ iff $\Box$for all objects x (x exemplifies Φ iff x exemplifies Ψ).

Thus, by what we've just said, we get that the two properties of being self-identical and being contained in one's singleton are necessarily equivalent.

Now, on the standard theory of properties in the possible worlds framework **IPT**, necessarily equivalent properties are identified. Remember that according to **IPT**, properties are *functions* and individuated as such. More specifically, we get:

Property Identity in IPT. For all properties Φ and Ψ, we have that $\Phi = \Psi$ iff for all worlds $w \in \mathcal{W}$, $\Phi(w) = \Psi(w)$.

An immediate consequence of **Property Identity in IPT** is that any two necessarily equivalent properties are *identified.* In particular, we get that the property of being self-identical simply *is* the property of being contained in one's singleton.

But, given what we said before, if we assume that an essential property of an object has to be part of the object's nature, then it is intuitively plausible to say that being self-identical is an essential property of Socrates, while being contained in {Socrates} is not. Certainly, both properties are classified by **PWA** as essential properties of Socrates. But intuitively, being self identical is part of the nature of Socrates, while being a member of {Socrates} is *not*: as we've said, intuitively, there is nothing in Socrates' nature that connects him to any set, and so also not to his singleton. But the two properties are actually *identical* according to **IPT**. Thus, assuming **IPT**, we simply *cannot* say that the one property is essential to Socrates while the other is not—they are one and the same property.

In philosophical jargon, we say that a context is *hyperintensional* iff in the context the substitution of necessary equivalents need not preserve truth-value [29, 103].[39] In contrast, a context is called *intensional* iff in the context the substitution of mere equivalents need not preserve truth-value, but the substitution of necessary equivalents always preserves truth-value.[40] In this terminology, we may identify the problem at hand as being that if we assume that the necessary properties of an object have to be part of the object's nature, then *ascriptions of essential properties create hyperintensional contexts*. This is nicely illustrated by the case of Socrates and his properties of being self-identical and being contained in his singleton: the properties of being self-identical and being contained in his singleton are necessarily equivalent, but the statement "being self-identical is an essential property of Socrates" is true, while the statement "being contained in his singleton is an essential property of Socrates" is false. In contrast, by what we've said above, the modal analysis **PWA** together with the intensional property theory **IPT** can only account for a notion of essential properties where

[39]Cresswell [29] defines hyperintensionality as the failure of the substitutivity of *logical* equivalents, but in recent times it has become common to adopt the above, weaker definition in terms of necessary equivalence [103]. Note that we speak somewhat imprecisely of *(necessary) equivalents*, without saying what kinds of things are supposed to be (necessarily) equivalent here. The reason for this intended imprecision is that we do not wish to repeat everything that we say about properties for propositions, sentences, and so on for all the things for which it makes sense to say that they are equivalent.

[40]Here, by *mere equivalents* we mean actual equivalents: two things that are actually equivalent. Traditionally, contexts where the substitution of mere equivalents always preserves truth-value are called *extensional contexts*. Thus, we may equivalently define an intensional context as a non-extensional context, where the substitutivity of necessary equivalents holds. Now, certainly, if the substitutivity of necessary equivalent fails, then also the substitutivity of mere equivalents must fail. For assume that the former fails and the latter does not. If we then take two necessary equivalents, they surely are also mere equivalents, since whatever is necessarily the case is also actually the case. Thus, they are substitutable—contradiction! As a consequence, we may alternatively define a hyperintensional context as one where neither substitution of mere equivalents nor the substitution of necessary equivalents need preserve truth-value: a hyperintensional context is a non-extensional and non-intensional context.

ascriptions of essential properties create intensional contexts—and according to Fine that is not enough.

More generally, Fine's point is that the framework of possible worlds is inherently ill-equipped to deal with concepts that create hyperintensional contexts. In the semantics that we've discussed in §1.1, the substitutivity of necessarily identical terms, necessarily equivalent formulas, and necessarily equivalent predicates all hold. For this reason, semantics in the possible worlds framework is usually called *intensional semantics* [50]. Fine writes:

> *Given the insensitivity of the concept of necessity to variations in source, it is hardly surprising that it is incapable of capturing a concept which is sensitive to such variation. Each object, or selection of objects, makes its own contribution to the totality of necessary truths; and one can hardly expect to determine from the totality itself what the different contributions were. One might, in this respect, compare the concept of necessity to the concept of communal belief, i.e. to the concept of what is believed by some member of a given community. It would clearly be absurd to attempt to recover what a given individual believes from what his community believes. But if I am right, there is a similar absurdity involved in attempting to recover the essential properties of things from the class of necessary truths.* [39, p. 9]

By what we just said, we can rephrase Fine's point in the following way: assuming that it's necessary for a property to be essential to an object, the property has to be part of the nature of the object, the concept of essential properties is hyperintensional—in the sense that ascriptions of essential properties create hyperintensional contexts—and the framework of possible worlds and necessity *de re* is, almost by definition, incapable of dealing with hyperintensional concepts.

Let's sum up the results of this section. We've seen that, given that we assume that it's necessary for a property to be essential to an object that the property is part of the nature of the object, Fine has given us a wide range of counterexamples to the modal analysis **PWA** supplemented with the intensional property theory **IPT**. More specifically, given the assumption that the essential properties of an object have to part of the object's nature, Fine has argued that weak necessity *de re* is not sufficient for essential properties. In other words, the following principle is false according to Fine:

Weak Necessity Implies Essence. For all objects x and all properties Φ, if at every possible worlds where x exists, x exemplifies Φ, then Φ is an essential property of x.

Moreover, Fine not only argues that the principle is false, he argues that the whole approach of analyzing essential properties in terms of necessity *de re* within the framework of possible worlds is fundamentally misguided. Fine *does* believe that weak necessity *de re* is necessary for essential properties (see footnote 30, p. 13). In other words, he subscribes to the following principle:

Essence Implies Weak Necessity. For all objects x and all properties Φ, if at every possible worlds where x exists, x exemplifies Φ, then Φ is an essential property of x.

But he argues that despite this connection between the concept of essential properties and the concept of weak necessity *de re*, we should not "get our hope up," as it were, about a workable modal analysis of essential properties. Since, given his assumption that it's necessary for a property to be essential to an object that the property is part of the nature of the object, ascriptions of essential properties create hyperintensional contexts and the framework of necessity *de re* is simply incapable of properly analyzing concepts that create hyperintensional contexts.

1.3 Grounding and Essential Properties

We shall now turn to a new analysis of essential properties in terms of *metaphysical ground*. We'll say more about metaphysical ground in due course, but for now, as a first approximation, we may think of the relation of (metaphysical) ground as the relation of one thing being the case *(wholly) in virtue of* a possible plurality of other things [42].[41]

We find the motivation for our new analysis in a certain analogy between defining the meaning of a term and giving the essence of an object proposed by Fine. Here we shall not go into the details of Fine's analogy and go straight to the heart of the matter. Fine writes:

> *We have seen that there exists a certain analogy between defining a term and giving the essence of an object; for the one results in a sentence which is true in virtue of the meaning of the term, while the other results in a proposition which is true in virtue of the identity of the object. However, I am inclined to think that the two cases are not merely parallel but are, at bottom, the same.* [39, p. 13]

[41]As Fine [42, p. 37–38] points out, there are different ways of expressing the relation of ground: we may say that one thing holds *because of* other things, where the "because" is read as *metaphysically because*, we may say that the some things are *metaphysically prior* to the thing, and so on. Here and in the following, however, we shall take "in virtue of" as our standard phrase for expressing ground.

By taking the phrase "in virtue of" to express metaphysical ground, we obtain at the

Essence Grounded in Identity (EGI). For all properties Φ and all objects x, Φ is an essential property of x iff x exemplifies Φ (wholly) in virtue of the identity of x.

It is unclear whether Fine would actually endorse this account, since in [42, p. 74–80] he goes to great lengths to argue that the concepts of essence and ground should be kept strictly apart. In [48], however, he revokes this earlier assessment and sketches a unified foundations for the two concepts. Here, we shall not go into the details of the relationship between ground and essence from a general perspective and we shall not try to figure out what exactly Fine's current position might be. Instead, we shall focus on the concrete view **EGI**, which we'll simply motivate by Fine's quote above.

As, we've said we'll have to say more about the nature of ground and its relata in due course, but for now, let's look at a few intuitive examples to illustrate the idea. Think of Socrates and Xanthippe again. Now, it seems plausible to say that Socrates exemplifies being a man in virtue of his identity: intuitively, it is part of being Socrates that he is a man. In contrast, it's similarly plausible that Socrates does not exemplify being married to Xanthippe in virtue of his identity: intuitively, there is nothing about Socrates identity in virtue of which he is married to Xanthippe. So, **EGI** fares well in light of our paradigmatic cases of essential properties. But more importantly, Fine's analysis can deal with the counterexamples to the modal analysis **PWA**. First, in the case of Socrates and his singleton it's plausible to say that Socrates is a member of {Socrates} in virtue of the identity of {Socrates} and *not* in virtue of the identity of Socrates. Thus, according to **EGI**, having Socrates as a member *is* an essential property of the singleton, but being a member of his singleton *is not* an essential property of Socrates. Similarly, it's plausible to say that Socrates is self-identical in virtue of his identity, but, intuitively, Socrates is not distinct from the Eiffel Tower *wholly* in virtue of his identity—also the Eiffel Tower's identity plays a role in this. And thus according to **EGI** being self-identical *is* an essential property of Socrates, while being distinct from the Eiffel Tower is not. Third, it's not plausible to say that Socrates exemplifies being such that there are infinitely many prime numbers in virtue of his identity—rather this is the case in virtue of the nature of the natural numbers. And so, according to **EGI**, being such that there are infinitely many prime numbers is *not* an essential property of Socrates—just as we want to say. In short, **EGI** agrees with our metaphysical assumptions not only in the paradigmatic cases, but also in the problematic cases of Fine's counterexamples.

Now, **EGI** is framed in terms of an informal use of the phrase "in virtue of," which we take as expressing the relation of metaphysical ground. But if we wish to take **EGI** as the basis for a reputable metaphysical analysis of essential properties—and we do—then we arguably need a proper philosophical background theory of metaphysical ground. A first step in this direction would be to properly regiment the use of the phrase "in virtue of" or, in other words, to give a *syntax of ground*. In the literature on ground, already the question how to formalize ground—what syntax of ground to choose—is a contentious issue [42, p. 46–48, 24, p. 253–54]. Indeed, we shall address this issue more prominently later in this dissertation. For the purpose of this introduction, however, let's simply focus on the expression "in virtue of" as our standard expression for ground.

Syntactically speaking, the phrase "in virtue of" falls into a somewhat strange category: in natural language, the expression takes *sentences* (or formulas) to the left and *terms* (or perhaps sequences of terms) to the right. Compare, for example, the expression "x exemplifies Φ in virtue of the nature of x." In this expression, the argument to the left of the phrase "in virtue of" is the open formula "x exemplifies Φ" and the argument to the right of the phrase is the term "the nature of x." But from a metaphysical perspective, this syntax suggests an undesirable view of ground: it suggests that ground is a relation that holds between what is expressed by sentences—propositions, facts, or the like—and what is denoted by terms—objects. But ground is standardly viewed as a relation between what is expressed by sentences (or perhaps formulas), i.e. facts, propositions, or the like [42].[42] For this reason, in the logic of ground, it is common to treat "in virtue of" simply as an operator, which takes sentences (or formulas) to both sides [42, p. 46]. Following this convention, at least for the purpose of this introduction, we'll take the logical form of a statement of ground to be:

$$\varphi \text{ in virtue of } \psi_1, \psi_2, \ldots,$$

where φ is a formula and $\psi_1, \psi_2, \ldots$ is a sequence of formulas. And the intended reading of a formula of the form φ in virtue of $\psi_1, \psi_2, \ldots$ is that what φ says is the case in virtue of what $\psi_1, \psi_2, \ldots$ say being the case.

Having regimented the use of the phrase "in virtue of," we wish to rephrase the analysis **EGI** in terms of this syntax. But now we face a problem: As we've said in natural language "in virtue of" takes terms to its right; and indeed in the analysis **EGI** we have the term "the identity of x" in this position. But we've agreed that we wish to view "in virtue of" as an operator, which takes formulas as arguments. So what formula should go into the place of "the identity of x" in the analysis **EGI**? We propose to take the formula "x being x" in its place. Metaphysically speaking, the idea is that instead of talking of the identity of a thing as an *object*, we view the identity of

[42] But compare [123], who holds that ground is a relation that can obtain between things from various different ontological categories.

an object as a *property*. Traditionally, the kind of property that we have in mind here is called a *haecceity*: the property of being a certain object. So, for example, the haecceity of Socrates is simply the property of being Socrates, the haecceity of Xanthippe is the property of being Xanthippe, and so on. And we express the haecceity of an object x simply by the predicate "x being x." Thus, applying this idea to the principle **EGI**, we get the following regimented version of the principle:

Essence Grounded in Haecceities (EGH). For all properties Φ and for all objects x, Φ is an essential property of x iff x exemplifies Φ in virtue of x being x.

In short, the essential properties of a thing are simply the properties that are grounded in its haecceity. We take it that this analysis is what Fine had in mind when he formulated **EGI**.

Given some standard assumptions in the logic of ground, **EGH** has some interesting consequences. First, it is standardly assumed that ground is *factive* in the sense that the relation can only hold between things that are the actually case [42, p. 48–50]. This assumption is more or less explicit in our manner of speaking: we have said that ground is the relation of one thing being the case in virtue of others. From a logical perspective, this assumption sanctions the following rules of inference for all formulas $\varphi, \psi_1, \psi_2, \ldots$:

$$\frac{\varphi \text{ in virtue of } \psi_1, \psi_2, \ldots}{\varphi}\ Fact_L \qquad \frac{\varphi \text{ in virtue of } \psi_1, \psi_2, \ldots}{\psi_i}\ Fact_R$$

Using these rules, we can see that according to **EGH**, objects exemplify all their essential properties. For assume that a property Φ is an essential property of an object x. Then it follows by **EGH** that x exemplifies Φ in virtue of x being x. But then, by a simple application of $Fact_L$, we can infer that x exemplifies Φ. This fact directly corresponds to the fact that according to the modal analysis **MA**, existing objects exemplify all their essential properties (see p. 2).

Second, it is usually assumed that *ground implies consequence*, in the sense that what φ says can only be the case in virtue of what $\psi_1, \psi_2, \ldots$ say being the case, if it's necessary that whenever what $\psi_1, \psi_2, \ldots$ say is the case, then what φ says is the case [121, p. 118, 42, p. 38].[43] Logically, this assumption sanctions the following rule of inference for all formulas $\varphi, \psi_1, \psi_2, \ldots$

$$\frac{\varphi \text{ in virtue of } \psi_1, \psi_2, \ldots}{\Box(\text{if } \bigwedge\{\psi_1, \psi_2, \ldots\}, \text{ then } \varphi)}\ Cons$$

Here, $\bigwedge\{\psi_1, \psi_2, \ldots\}$ simply denotes the (possibly infinitary) conjunction of $\psi_1, \psi_2, \ldots$. Thus, it follows from **EGH** that it's necessary *de re* for an object

[43] But see [128] for a disagreeing argument.

that if it exemplifies its haecceity, then the object exemplifies its essential properties. For assume that a property Φ is an essential property of an object x. Then it follows by **EGH** that x exemplifies Φ in virtue of x being x. And from this we can infer by a simple application of *Cons* that $\Box$(if x exemplifies being x, then x exemplifies Φ).

Now, since we have assumed that an object can only exemplify a property if the object exist (see p. 4), we get that an object can only exemplify the property of being that object if the object exists. Conversely, if an object exists, then it clearly intuitively exemplifies the property of being that object. Moreover, both of these principles are plausibly regarded as laws of metaphysics and thus they are metaphysically necessary. Putting both of this together, we get the following modal connection between haecceities and existence:

Haecceities and Existence. For all objects x, $\Box$(x exemplifies being x iff x exists).

Together with the previous observation, namely that according to **EGH**, if a property Φ is an essential property of an object x, then $\Box$(if x exemplifies being x, then x exemplifies Φ), we get by **Haecceities and Existence** that if Φ is an essential property of x, then $\Box$(if x exists, then x exemplifies Φ). Semantically speaking, if Φ is an essential property of x, then in every possible world where x exists, x exemplifies Φ. In other words, **EGH** entails **Essence Implies Weak Necessity**. Thus, just like on the modal analysis **MA**, if a property is contingent to an object—if it's possible *de re* for the object to exist without exemplifying the property—then the property is not an essential property of the object (see p. 3).

Conversely, however, **EGH** does *not* entail the principle **Weak Necessity Implies Essence**. Indeed, as illustrated by Fine's counterexamples, as we've discussed above, the principle **Weak Necessity Implies Essence** is *false* according to **EGH**: there are weak necessities *de re* of an object that are not essential properties of the object. More generally, it is usually assumed that the relation of ground requires relevance between the *grounds*—the things expressed to the right of a true "in virtue of"-statement—and the *groundee*—the thing that is expressed to the left of a true "in virtue of"-statement. Thus, for statements $\varphi, \psi_1, \psi_2, \ldots$ and θ, the rule of inference

$$\frac{\varphi \text{ in virtue of } \psi_1, \psi_2, \ldots}{\varphi \text{ in virtue of } \theta, \psi_1, \psi_2, \ldots}\ Weakening$$

is not *sound*: it can lead from a true premise to a false conclusion. Thus, there is a natural sense in which **EGH** satisfies the **Relevance Condition**: If we say that what φ says is (wholly) relevant to an object x iff what φ says is grounded in x being x, it follows that if Φ is an essential property of an object x, then that x exemplifies Φ is relevant to x—a reasonable way to say that the property Φ is relevant to x. Moreover, the idea that an object having a property in virtue of the object being that object is a reasonable gloss of the property being part of the object's nature. Thus, in this sense, **EGH** satisfies Fine's condition that it's necessary for a property Φ to be an essential property of an object x, Φ has to be part of the nature of x.

Finally, metaphysical ground is, intuitively, *hyperintensional*, in the sense that in the context of "in virtue of" the substitution of necessary equivalents need not preserve truth-value. This is readily illustrated by the example of Socrates being self-identical and Socrates being a member of his singleton. As we've said above, the two properties of being self-identical and being the member are necessarily equivalent. But the sentence "Socrates is self-identical in virtue of being Socrates" is intuitively true, while the sentence "Socrates is a member of his singleton in virtue of being Socrates" is intuitively false. Thus, the concept of ground has an important property that necessity *de re* lacks: it creates hyperintensional contexts. For the logic of ground this means that in contexts of "in virtue of" the substitutivity of equivalents fails. More specifically, for formulas $\varphi, \psi_1, \psi_2, \ldots$, and θ, the rules of inference

$$\frac{\varphi \text{ in virtue of } \psi_1, \psi_2, \ldots \quad \Box(\varphi \text{ iff } \theta)}{\theta \text{ in virtue of } \psi_1, \psi_2, \ldots}\ Subs_L$$

$$\frac{\varphi \text{ in virtue of } \psi_1, \psi_2, , \ldots, \psi_i, \ldots \quad \Box(\psi_i \text{ iff } \theta)}{\varphi \text{ in virtue of } \psi_1, \psi_2, \ldots, \theta, \ldots}\ Subs_R$$

are not *sound*: they can lead from true premises to false conclusions. Thus, we have to be careful when reasoning with "in virtue of." But this property of ground also has its upsides: it is effectively this property—that ground creates hyperintensional contexts—which allows **EGH** to deal with Fine's counterexamples. Above we have argued that under theassumption that the essential properties of an object have to be part of its nature, ascriptions of essential properties create hyperintensional contexts (see p. 24). Thus, we *need* a hyperintensional concept, if we wish to analyze such a notion of essential properties. And indeed, by what we have said above, ground appears to fit the bill: with the help of "in virtue of" we can distinguish between Socrates being self-identical and Socrates being a member of his singleton. Thus, all in all, **EGH** fares quite well with respect to our essentialist intuitions and the conditions that arose from the discussion of Fine's counterexamples.

But, for Fine's analysis **EGH** to become a respectable analysis of essential properties, which is metaphysically on *a par* with the modal analysis, two essential ingredients are missing: First, we need a proper background theory of metaphysical ground, which can play the same role that the possible worlds framework plays for the analysis: by translating **MA** through

Kripke's semantics into the possible worlds analysis **PWA**, modal essentialists were able to disperse Quine's skeptical worries. Now, a skeptic about ground might raise a similar worry about **EGH** as Quine did about **MA**: they might argue that the notion of metaphysical ground is conceptually confused and thus the analysis **EGH** is confused as well. To (perhaps pre-emptively) counter this objection, we had better come up with a suitable semantics for the concept of ground, comparable in scope and fruitfulness to the possible worlds semantics for necessity *de re*. Second, as we've said above, part of the reason why **PWA** was so successful as an analysis of the concept of essential properties was that it could be supplemented with a working theory of properties in the same framework that the analysis was formulated in: the theory intensional property theory **IPT** (see p. 11). We know already that **IPT** will not work as a background theory of properties for **EGH**: this is illustrated by the example of Socrates being self-identical and Socrates being a member of his singleton; both properties are necessarily equivalent, and so identical according to **IPT**, but, according to **EGH**, being self-identical *is* and essential property of Socrates, while being contained in his singleton is *not*. So, in addition to a semantic framework for ground, we need a property theory to get **EGH** "off the ground," as it were. And this property theory will need to be able to distinguish between necessarily equivalent properties: it needs to be a *hyperintensional* property theory.

Now, at the present stage of research, these two components are still missing: A lot of exciting research has recently been carried out in the logic and semantics of ground [21, 133, 15]. But semantic theories of ground are a relatively young field of research, and we are not at the level of refinement of Kripke's possible worlds semantics for necessity *de re*. One thing that is notoriously missing is a semantic treatment of occurrences of "in virtue of" in the context of other occurrences of "in virtue of" or, as it's sometimes called, occurrences of *iterated ground*. Litland [88, p. 131–78] and Litland [87] gives a proof system that deals with iterated ground, but a semantic treatment is still missing. Philosophically speaking, however, if we wish **EGH** to succeed as an analysis of the concept of essential properties on a par with **MA** or **PWA**, we should like to have a semantic treatment of "in virtue of" that can accommodate iterated ground. To illustrate, consider an object x and a property Φ such that Φ is an essential property of x. Now we may ask ourselves if it is also an essential property of x that Φ is an essential property of x. According to **PWA**, this question has a clear answer: If in every worlds where x exists, x exemplifies Φ, then every world is such that in every world where x exists, x exemplifies Φ—in other words, it is an essential property of x that Φ is an essential property of x. According to **EGH**, in contrast, the answer involves a case of iterated ground. According to **EGH**, it is an essential property of x that Φ is an essential property of x iff that Φ is an essential property of x is grounded in the haeccaeity of x. But if we understand Φ being an essential property of x as x exemplifying Φ being grounded in the haecceity of x, then we get that it is an essential property of x that Φ is an essential property of x iff the fact that x exemplifies Φ in virtue of its haecceity is in turn again grounded in the haeccaeity of x.[44] But this is a question about iterated ground! Thus, if we wish to be able to tell whether it is essential to objects that they have their essential properties—and, of course, we do—we need to be able to account for iterated ground. In other words, once we adopt **EGH** as our metaphysical analysis of the concept of essential properties, we're naturally lead to questions of iterated ground. And consequently, if we wish **EGH** to succeed as a metaphysical analysis of essential concepts, we need a semantic account of iterated ground.

With regard to suitable property theories, the situation is even worse. There are hyperintensional theories of properties on the market [8, 9, 130], but none of these theories is formulated in a framework that is suitable for the semantic analysis of the concept of ground. But, as we've pointed out, if we wish **EGH** to succeed as a metaphysical analysis, we need such a theory. Thus, the two central aims of this dissertation will be to provide these two missing pieces: (1) a semantic analysis of ground that can accommodate cases of iterated ground and (2) a hyperintensional property theory in a framework suitable to a semantic analysis of ground. The core chapters of this dissertation will deal with exactly these two issues.

1.4 The Problem of Explicating Essential Properties

Let's take stock. So far, we've discussed two analyses of essential properties: the modal analysis in the form of **PWA** supplemented with the intensional

[44]There is a delicate (logical) issue here: In formulating **EGH**, we've followed the convention of formulating an analysis as an *equivalence*. Since **EGH** could reasonably be considered a principle metaphysics, it this equivalence is even metaphysically necessary. But, as we've pointed out above, in contexts of "in virtue of" the substitutivity of equivalents fails. In particular, **EGH** does not warrant us to freely replace "Φ is an essential property of x" with "x exemplifies Φ in virtue of the haecceity of x" in contexts of ground. However, the idea behind **EGH** certainly is that the two formulas *mean* the same: the idea is that a property being essential to an object *means* that the object's exemplifying the property is grounded in its haecceity. And intuitively, synonymous formulas should be substitutable even in ground-theoretic contexts. Note that, on this concept of meaning, necessary equivalence is *not* sufficient for synonymy—the concept is hyperintensional. The bottom line is that to justify the inference we would need a stronger principle than **EGH**. In ground-theoretical contexts a definition should indeed be such that the definiens and the definiendum are substitutable for one another. But developing a corresponding notion of definitions is a non-trivial issue. We'll address the issue later in this dissertation (see chapter 2, §**??**, p. **??**-**??**). For now, we use the informal reasoning in this footnote to justify the inference in the main text above.

property theory **IPT** and the ground-theoretic analysis **EGH**. We've seen that the even though **EGH** still lacks a proper semantic formulation and a background property theory, the analysis can deal with all of Fine's counterexamples to the modal analysis and it is intuitively well motivated. We might ask at this point: But which analysis is the *correct* analysis of the concept of essential properties? We wish to argue that none of them is—but not for the reasons one might perhaps think. We wish to argue that there is no *one correct* analysis of the concept of essential properties.

Remember that the crucial issue in the case of Fine's counterexamples was that we assume that in order for a property to be essential to an object, the property has to be part of the object's nature. But why should we share this assumption? Fine does not offer an *argument* for his assumption, but rather a *plausibility* claim. He writes:

> *I am aware, though, that there may be readers who are so in the grip of the modal account of essence that they are incapable of understanding the concept in any other way. One cannot, of course, argue a conceptually blind person into recognizing a conceptual distinction, any more than one can argue a colour blind person into recognizing a colour distinction. But it may help such a reader to reflect on the difference between saying that singleton Socrates essentially contains Socrates and saying that Socrates essentially belongs to singleton Socrates. For can we not recognize a sense of nature, or of 'what an object is', according to which it lies in the nature of the singleton to have Socrates as a member even though it does not lie in the nature of Socrates to belong to the singleton?* [39, p. 5]

We're inclined to agree: One cannot argue for one concept being the one correct concept of essential properties. There is a concept of essential properties according to which it's necessary for a property to be essential to an object that the property is part of the object's nature. But there is also a concept according to which this is not necessary: there are (at least) *two* concepts of essential properties. Both concepts are coherent and neither has a claim to being the correct one.

We can think of the concept that modal metaphysicians are talking about as defined **MA**, viewed as a definition of essential properties in terms of necessity *de re*. This definition is illustrated by means of paradigmatic examples: being a man, for example, is supposed to be an essential property of Socrates according to **MA**, while being married to Xanthippe is not. The concept of necessity *de re* involved in **MA** is furthermore regimented by means of the standard modal logic, which allows us to ascertain the consequences and applications of **MA**. Finally, the (more or less) informal definition **MA** is made precise in the framework of possible worlds by means of the analysis

PWA supplied with the intensional property theory **IPT**. Now, as we've seen, the concept of the nature of an object in Fine's sense is not part of the framework of possible worlds. The concept that Fine has in mind is hyperintensional: the two properties of being self-identical and being a member of one's singleton are necessarily equivalent, yet the one is part of the nature of Socrates in Fine's sense, while the other is not. But according to **IPT** this claim cannot even be coherently formulated, since according to the theory the two properties are identical. Moreover, hyperintensional concepts can in general not be analyzed in the *intensional* framework of possible worlds. In other words, Fine's concept of the nature of an object simply isn't part of the framework of possible worlds. And therefore, Fine's counterexamples don't really arise in the framework and in particular they don't show that the analysis is incorrect or incoherent.[45]

But also the concept of essential properties as properties grounded in the identity of things makes intuitive sense. We can think of this concept as defined by **EGH**, viewed as a definition of essential properties in terms of ground. Also this definition is illustrated by means of examples: the property of being self-identical, for example, is supposed to be an essential property of Socrates according to **EGH**, while being contained in {Socrates} is not. The use of the phrase "in virtue of" in **EGH** is furthermore regimented by the standard principles in the logic of ground, which also allow us to ascertain the consequences of the definition. Thus, **EGH** is, at least pre-theoretically, precise. Fine's examples show that the modal analysis and the ground-theoretic analysis don't agree, not even pre-theoretically: according to **MA**, being a member of his singleton is an essential property of Socrates, while according to **EGH** it is not. The underlying reason for this discrepancy is that in the ground-theoretic framework, the notion of the nature of an object can be formulated and indeed it is required for a property to be essential to an object that the property is part of the nature of the object—in just the sense we discussed above. The only thing that's missing is a semantic precisification in the form of a semantic analysis of "in virtue of" and a fitting property theory to go along. And this is exactly what we wish to do in this dissertation.

The point here is that we don't have *two* competing analyses for *one single* concept of essential properties, but, at least once we've supplied the missing

[45]Some authors, like [139, 135], respond to Fine's counterexamples from the perspective of the modal analysis, aiming to vindicate the analysis. But both these authors abandon the "pure" framework of possible worlds in favor of some extended framework that can account for hyperintensional distinctions: Zalta [139] moves to the framework of his *object theory*, and Wildman [135] uses the concept of *sparse properties*. Both these extensions allow the authors to make hyperintensional distinctions. In this dissertation, however, the modal account is not the prime focus, we'll talk mainly about the ground-theoretic approach given by **EGH**.

components to the ground-theoretic concept, we really have *two* analyses of *two different* concepts of essential properties. So which analysis should we adopt? It appears that the most important criterion is given by what we want to do with the concept of essential properties and its analysis. And there appear to be interesting applications for both concepts and their analyses: The modal analysis allows us to say what is metaphysically necessary for an object to exist, or, if we may put this way, what is essential to the object's existence. The ground-theoretic analysis, in contrast, allows us to say what is determined by an object being that very object, or, if we may put it this way, what is essential to the object's haecceity. These are two different questions and so we should expect two different approaches and two different answers. We take it that Fine's central contribution to the issue is to raise awareness for the second question. This question is, of course, not a new one, but, at least until recently, it was largely ignored in essentialist metaphysics. As Fine points out, both analyses—the modal and the ground-theoretic one—ultimately trace back (at least) to Aristotle [39, p. 2–3]. But at least among analytic metaphysicians the modal question took precedence for the second half of the last century. Following Fine's paper, however, the ground-theoretic question took center stage. And there's work to be done—here lies the area of research that this dissertation aims to contribute to.

It might be helpful to phrase the point of the section in a slightly different way: We can view the problem of analyzing the concept(s) of essential properties as a Carnapian *problem of explication* [17, p. 1–18]. According to Carnap: "the task of *explication* consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second" [17, p. 3, emphasis in the original]. In a given explication, Carnap calls the explicated concept the *explicandum*, and the explicating concept the *explicatum*. Before giving an explication of a concept, so Carnap, we should try to make the explicandum as precise as possible: for example, we should give an informal definition of the explicandum, we should give examples of the correct application of the concept, or we should discuss its logic [17, p. 4–5]. Only then, we should attempt to find a proper explicatum. Once we've determined an explicatum, according to Carnap, the quality of the explication can be judged by four criteria: (1) the *similarity* between the explicatum and the explicandum; (2) the *exactness* of the explicatum; (3) the *fruitfulness* of the explicatum; and (4) the *simplicity* of the explicatum [17, p. 5–8].

Now, given what we've said in this section, we can naturally regard the dialectic about the modal analysis of essential properties as an explication of the first concept of essential properties: Here, the explicandum is the concept of a property being an essential to an object, where we take this to be the case iff it's necessary *de re* for the object to exist that it exemplifies the property. The analysis **MA** gives us a more or less informal definition of this explicandum. The intuitions that being a man is an essential property of Socrates in the relevant sense and that being married to Xanthippe is not give us paradigmatic examples for the correct application of the concept. And with the use of philosophical modal logic, we can ascertain the basic logic of the concept. Then, **PWA** gives us a proper the explication of the concept: a property being essential to an object (in the relevant sense) is explicated as the object exemplifying the property in every world where it exists—our explicatum. This explication is then made more precise by supplying it with the intensional property theory **IPT**.[46]

And viewed as an explication, **PWA** fares quite well with regard to Carnap's four criteria. As we've seen, given our assumptions about possible worlds, **PWA** and **MA** agree on the paradigmatic examples. Indeed, by Kripke's soundness and completeness result, the modal logic that we've used to determine the informal logic of the concept of essential properties defined by **MA** is semantically mirrored by the concept given by **PWA**. So **PWA** fares quite well with regards to the criterion of similarity (1). Moreover, Kripke's semantic framework of possible worlds, especially when supplemented with **IPT**, is formally exact. In particular, all the relevant concepts involved in **PWA** are well-defined in this framework. Indeed, the framework is precise enough so that we can use standard logical methods to establish the soundness and completeness result mentioned before. Thus, criterion (2) is also satisfied to a high degree. As we've said the framework of possible worlds is in itself fruitful: it allows us to analyze a wide range of philosophical concepts. Thus, the explication **PWA** is fruitful as well, as it allows us to connect the relevant concept of essential properties with these other concepts. This makes it plausible that the criterion of fruitfulness (3) is well satisfied. And finally, the explication **PWA** is arguably simple: all the concepts that are involved are existence and exemplification at (all) possible worlds. Hence, the criterion of simplicity (4) is arguably well satisfied.

Now, the concept of a property being essential to an object iff the object exemplifying the property is grounded in the object's haecceity poses a *different* problem of explication. In this introduction, we've taken care the first step of this project: we've tried to make the explicandum as precise as possible. The analysis **EGH** gives us an informal definition of the relevant concept of essential properties. Fine's examples give us a good indicator for the correct application of the concept. And the logic of ground gives us a good idea of the logic of the concept. Now it is time for the next step: to give an explication of the concept of essential properties defined by **EGH**. In this dissertation, we wish to tackle this project by providing two things: (1) a

[46] Already Kripke's semantics can be regarded as an explication of the concept of necessity *de re*, but here we focus on the concept of essential properties, which is defined in terms of necessity *de re*.

semantic analysis of the concept of metaphysical ground which can account for iterated ground, and (2) a hyperintensional property theory in the same semantic framework. Our hope is that by providing these two things, we can bring the concept given by **EGH** to a comparable level of precision as the concept given by **MA** achieves through **PWA** and **IPT** in the semantic framework of possible worlds. But, since we view the problem at hand as a problem of explication, ultimately its success will have to be judged by means of Carnap's four criteria.

1.5 Overview of the Thesis

To conclude this introduction, let's sum up the concrete goals that we've come up with. In this dissertation we wish to supply the ground-theoretic analysis of essential properties **EGH** with the following two components:

(1) a semantic analysis of metaphysical ground, which in particular can account for cases of iterated ground and can play the same role that the possible worlds framework plays for **MA**, effectively giving us an analog to **PWA**, and

(2) a hyperintensional property theory in the same semantic framework as our semantic analysis, which can play the same role that **IPT** plays for the **PWA** in the framework of possible worlds.

By providing these two components, we aim to give an explication of the concept defined by **EGH** in the Carnapian sense. Thus, as the standards for success of our project we take the four criteria listed by Carnap: similarity, (formal) precision, fruitfulness, and simplicity.

We will pursue those two goals over the course of the core chapters of this dissertation. In chapters two and three, we approach the first of the two goals. There is a fundamental distinction among theories of ground according to what they take to be the relata of ground, or rather how these relata are individuated.[47] *Conceptualist* theories of ground take the relata of ground to be conceptually individuated *truths*: fact-like entities that are individuated by the sentences or propositions that express them. *Worldly* theories,

[47]Here we only talk about theories of ground that take ground to be a relation between fact-like entities. This was already implicit in the way we informally introduced ground as the relation of one thing *holding* in virtue of others—only fact-like entities can be coherently said to *hold*. There are other kinds of theories of ground, which allow for ground to relate things from different ontological categories, see, for example, [123]. We prefer to keep the two concepts apart: ground holds between fact-like entities and a ground-like relation that can hold among things from different ontological categories is better conceived of as a form of *ontological dependence*. We'll not argue that point here in more detail, but in this dissertation, we'll only talk about ground in the previous sense.

in contrast, take the relata to be worldly individuated *facts*: facts that are individuated by the objects, properties, and relations they concern.[48] Now, in chapter two, we'll discuss a new approach to conceptualist theories of ground, where we formalize ground by means of a relation predicate of sentences. We'll develop different axiomatic theories of ground in this framework and show their consistency. However, we'll show that for the present purpose, this approach is not developed enough: in particular, if we try to deal with cases of iterated ground in this framework, we run into serious difficulties. Therefore, we'll abandon these theories for the purpose of this dissertation. Nevertheless, we argue that these theories present a new and exciting approach to theories of ground, which hopefully prove to be fruitful in future ground-theoretic research.

In chapter three, we will turn to worldly theories of ground. These theories are more developed than conceptualist theories of ground at the present stage of research. In particular, Correia [24] gives a treatment of worldly ground, which consists of a semantics together with a sound and complete proof system. Fine [42] gives an equivalent semantics in terms of truthmakers. However, neither Correia nor Fine treat cases of iterated ground in their semantics. In chapter three, we will extend Fine's version of the semantics to account for iterated cases of ground an we will give a proof theory the resulting semantics. The framework that we'll develop (or rather refine) in this chapter will be the basis for our explication of the concept defined by **EGH**. Thus, the view that we will be explicating in this thesis may be called *worldly ground-theoretic essentialism*: the view that at least some objects have at least some essential properties, viewed as properties that are grounded in their haecceities on a worldly conception of ground.[49]

Then, in chapter four, we will tackle the second aim of the thesis. In this chapter, we will propose a new hyperintensional property theory in the semantic framework of truthmaker semantics. We will argue that the other approaches to hyperintensional property theories on the market do not satisfy certain intuitive desiderata for hyperintensional property theories, and, as a consequence, they are not suited to the purpose of this dissertation. This chapter will be the last core chapter of the dissertation, and at its end, we will have arrived at the framework we desire: a semantic framework, complete with a theory of properties, which can deal with (iterated and uniterated cases of) metaphysical ground.

Finally, in the conclusion, we'll formulate an explication of the concept given by **EGH** in the framework that we've developed in the previous chapters.

[48] For more on different conceptions of facts and the particular terminology that we use here, see [41].

[49] *Conceptualist ground-theoretic essentialism* is also an enticing view, but it falls outside the scope of this thesis.

This will be the final result of this dissertation. We'll subject the explication to Carnap's four criteria. But ultimately, the success of the explication will have to be judged in the bigger scheme of things: how well the framework works in the bigger context of ground-theoretic metaphysics—the paradigm that metaphysics should be carried out with the help of ground-theoretic concepts. And this assessment goes (well) beyond the scope of this dissertation.

Chapter 2

Axiomatic Theories of Ground

2.1 Preface

This chapter contains a theory of ground that we're ultimately going to discard for the purposes of the dissertation. But the reasons for this are quite substantial and they will emerge only in the course of the chapter. The approach we take in this chapter is to formalize ground by means of a *relational predicate*, rather than by means of a *sentential operator*, as we will do in the rest of the dissertation. We'll argue that the predicate approach is both natural and philosophically well-motivated (see §2.3). Moreover, the approach allows us to obtain nice results connecting theories of truth and theories of ground (see §§2.5–2.6). But the approach also has problems. As we will show, we get *paradoxes of self-reference* (see §**??**) and we can't straightforwardly deal with the notion of *full* ground (see §2.10). Especially the latter point is problematic for us, since we wish to model essential properties as properties that are *fully* grounded in the identity of things. Ultimately, this will lead us to discard the approach—at least for the purpose of the dissertation.

But the chapter still has some merit. First, it is non-trivial to see that these problems arise. In particular the paradoxes of self-reference in the context of ground have not been discussed in the literature of ground so far.[1] Second, the results connecting theories of truth and theories of ground (§§2.5–2.6) are philosophically interesting in their own right. In particular, as we shall argue (§§**??** and 2.10), they motivate the application of methods and results from

[1] The only exception is my paper [72], which is attached as an appendix to this dissertation. See Appendix A.

theories of truth to the theory of ground. Thus, this chapter is ultimately a plea for collaboration between logicians working on theories of truth and metaphysicians working on theories of ground.

2.2 Introduction

Partial ground is the relation of one truth holding either wholly or partially in virtue of another [42, 44].[2] To illustrate the concept, consider a couple of paradigmatic examples:

(1) The truth of the disjunction that $5 + 7 = 12$ or $1 = 2$ holds wholly in virtue of the truth of its only true disjunct that $5 + 7 = 12$.

(2) The truth of the conjunction that $5+7 = 12$ and $2\times 2 = 4$ holds partially in virtue of the truth of its first conjunct that $5 + 7 = 12$ and partially in virtue of the truth of its other conjunct that $2 \times 2 = 4$.

Partial ground in this sense is a *strict partial order* on the truths: it is *irreflexive*—no truth partially grounds itself—and it is *transitive*—partial grounds are inherited through partial grounding.[3] Thus, partial ground gives rise to a *hierarchy of grounds*, in which the partial grounds of a truth rank "strictly below" the truth itself. The aim of this paper is to axiomatize this hierarchy over the truths of arithmetic.[4]

The main novelty of the paper is that we will use a ground *predicate* rather than an operator to formalize partial ground. This approach to formalizing partial ground has several philosophical benefits, which we will outline in

[2] For (opinionated) introductions to the concept(s) of ground, see [28, 42]. For an overview of the recent literature, see [15, 21, 118, 133]. Most research focuses on the notion of *full* ground: the relation of one thing holding wholly in virtue of a possibly plurality of other truths [42, p. 37]. For reasons that we will discuss more comprehensively in §2.10, we will focus on the notion of partial ground in this paper. For more on the distinction between full and partial ground, see [42, p. 50].

[3] This is, in any case, the standard view of partial ground. Some authors have challenged this view: Jenkins [66] challenges the claim that partial ground is irreflexive and Schaffer [122] challenges the claim that partial ground is transitive. See Litland [86] and Raven [119] for a defense of the standard view against these challenges.

[4] The main reason for taking arithmetic as the starting point here is that the standard theory of arithmetic, Peano arithmetic PA, can double in well-known ways as a theory of arithmetic *and* a theory of syntax (see §2.4). Thus, by taking PA as our starting point, we can effectively kill two birds with one stone: PA can function as the theory that tells us which sentences are true *and* function as a theory of syntax that allows us to talk about these sentences. Regardless of this technical convenience, nothing philosophically "deep" hinges on this particular theory choice. Note, however, that we're explicitly *not* including truths about partial ground in the hierarchy. There are specific technical and philosophical issues that arise in the context of such truths, which shall be discussed in the second part of the paper. See also our discussion of the issue on p. 52 of this article.

more detail in the following section. So far, however, most authors have eschewed the approach for reasons we'll discuss in detail in the following section as well. Ultimately, we argue, the benefits of the approach outweigh it's perceived drawbacks. Most importantly, the predicate approach will allow us to connect theories of partial ground with axiomatic theories of truth. In particular, once we've formulated the usually accepted principles of partial ground using a ground predicate, we can bring out the truth-theoretic commitments of theories of partial ground, in the sense that we can show that the resulting theory of partial ground is a *conservative extension* of the well-known theory PT of *positive truth* [58, p. 116–22].

2.3 The Predicate Treatment of Partial Ground

In this paper, we will formalize partial ground using the relational predicate $\lhd$ of sentences—our ground predicate. We'll add this predicate to the language of PA, where we may obtain a unique name $\ulcorner\varphi\urcorner$ for every sentence φ using the technique of Gödel-numbering. Here and in the following, we shall take the relata of partial ground to be (true) sentences, the idea being that partial ground is a relation on the *truths* (for further discussion of this assumption, see p. 47 below). Thus, we can formalize example (1) from above by:

$$\ulcorner\overline{5}+\overline{7}=\overline{12}\urcorner \ \lhd \ \ulcorner\overline{5}+\overline{7}=\overline{12} \ \vee \ \overline{1}=\overline{2}\urcorner,$$

where $\overline{n}$ is the numeral for the natural number n. In contrast, most authors formalize partial ground using the operator $\prec$ of sentences—the *(partial) ground operator*.[5] In the case of our example, these authors would add the ground operator to the language of PA, and then formalize example (1) by:

$$\overline{5}+\overline{7}=\overline{12} \prec (\overline{5}+\overline{7}=\overline{12} \ \vee \ \overline{1}=\overline{2}).$$

The syntactic difference between the two approaches is that the ground predicate takes terms denoting sentences as arguments, while the ground operator takes sentences themselves as arguments.

The predicational theory of partial ground that we will develop in this paper subsumes the standard operational theory of partial ground, in the sense that for all sentences φ and ψ, if $\varphi \prec \psi$ is derivable in the latter theory, then $\ulcorner\varphi\urcorner \lhd \ulcorner\psi\urcorner$ is derivable in our theory. The converse direction, however, does not hold in general: there are sentences φ and ψ such that $\ulcorner\varphi\urcorner \lhd \ulcorner\psi\urcorner$ is derivable in our theory, while $\varphi \prec \psi$ is *not* derivable in the standard theory of partial ground.[6] Thus, on the predicate approach we are able to obtain a strictly stronger theory of partial ground.

[5] Cf. [24, 43, 42, 44, 73, 88, 85, 121, 124]. Different authors may use different symbols for the ground operator.

[6] We will show this rigorously in §2.5.2.

But there are other reasons to prefer the predicate approach over the operator approach:

2.1 **Quantification over Truths**: The predicate approach has greater expressive strength than the operator approach. In particular, using the ground predicate, we can formalize ground-theoretic principles involving quantification over truths in a natural way. Take the two principles stating that partial ground is an irreflexive and transitive relation on the truths as an example. On the predicate approach, we can directly formalize these principles as:

$$(\text{Irreflexivity}_{\lhd}): \quad \forall x \neg (x \lhd x)$$

$$(\text{Transitivity}_{\lhd}): \quad \forall x \forall y \forall z (x \lhd y \land y \lhd z \to x \lhd z)$$

where the intended range of the quantifiers is the set of all truths.[7] On the operator approach, in contrast, we can (prima facie) only formalize these principles by affirming the instances of the following *schemata* for all sentences φ, ψ, and θ:

$$(\text{Irreflexivity}_{\prec}): \quad \neg(\varphi \prec \varphi)$$

$$(\text{Transitivity}_{\prec}): \quad (\varphi \prec \psi) \land (\psi \prec \theta) \to (\varphi \prec \theta)$$

Thus, on the operator approach, we can achieve quantification over truths only by moving to quantification over sentences in the *meta-language*, while on the predicate approach, we can directly express quantification over truths in the *object language*.[8]

Moreover, the strategy of moving to quantification in the meta-language fails once we consider principles involving existential quantifiers. Think for example of the intuitively plausible principle that a sentence is true iff its truth is either fundamental or grounded in some other truth. On the predicate approach, we can straightforwardly formalize this principle

[7] In the literature on ground, we distinguish between *factive* and *non-factive* conceptions of ground [cf. 42, p. 48–50]. On a factive conception, ground can only obtain between factive things, such as truths or facts. On a non-factive conception, the relation of ground can also hold between non-factive things, such as falsehoods or non-obtaining states of affairs. The notion of partial ground that we are working with in this paper is a factive notion of ground. Later we shall enforce this by means of axioms stipulating that the relation of partial ground can only hold between truths.

[8] A remark is in order: We could, of course, achieve similar results on the operator approach using quantification into sentence position or *propositional quantification*. But propositional quantification means a significant deviation from classical logic, while on the present approach we can comfortably stay within the purview of classical (first-order) logic. This highlights another benefit of the predicate approach: it allows us to study partial ground using entirely standard methods, well-known from first-order logic and model-theory.

as:

$$\forall x(Tr(x) \leftrightarrow (Fund(x) \vee \exists y(y \lhd x))),$$

where Tr is a unary *truth predicate* that applies to all and only the true sentences and $Fund$ is a unary predicate that applies to all and only the sentences whose truth is fundamental. Moreover, we could plausibly *define* this predicate $Fund$ by postulating that:

$$\forall x(Fund(x) \leftrightarrow_{def} Tr(x) \wedge \neg\exists y(y \lhd x)).$$

Then, in a predicational theory of ground with this definition, we'll be able to *derive* the claim that a sentence is true iff its truth is either fundamental or grounded in some other truth. On the operator approach, in contrast, we could not even formalize the principle in the first place: there simply is no way to express the nested universal and existential quantification over truths on that approach.

Finally, using quantification over truths, we can define useful ground-theoretic concepts directly in our object language. Take the concept of *weak* partial ground as an example [42, p. 51–53]. This is the relation of one truth being a "stand-in" for another in the context of partial ground. Following [42, p. 52], we can define weak partial ground in terms of our ordinary, *strict* notion of partial ground by saying that the truth of φ *weakly partially grounds* the truth of ψ just in case the truth of φ strictly partially grounds any truth that the truth of ψ grounds. It then follows, for example, that any truth weakly partially grounds itself, since clearly it strictly partially grounds any truth that it itself strictly partially grounds. Or, for another example, if the truth of φ strictly partially grounds the truth of ψ, then the truth of φ also weakly partially grounds the truth of ψ. This follows immediately from the transitivity of (ordinary strict) partial ground.[9] But conversely, it may very well happen that the truth of φ weakly partially grounds the truth of some ψ without strictly grounding it. Just think of the case where ψ is identical to φ: we've just seen that the truth of φ weakly grounds itself, but by the irreflexivity of strict partial ground (see p. 2) φ does not strictly partially ground itself.[10]

On the predicate approach, we can define a binary predicate $\trianglelefteq$ for this relation by postulating:

$$\forall x\forall y(x \trianglelefteq y \leftrightarrow_{def} \forall z(y \lhd z \rightarrow x \lhd z)).$$

[9] To see this, suppose that the truth of φ strictly partially grounds the truth of ψ and that the truth of ψ strictly partially grounds the truth of some arbitrary θ. It follows immediately by the transitivity of strict partial ground that the truth of φ strictly partially grounds the truth of θ, establishing that the truth of φ weakly partially grounds the truth of ψ.

[10] For a (critical) discussion of the concept of weak ground, see [33].

On the operator approach, in contrast, we can't define weak partial ground in this way—there we need to introduce a primitive operator $\preceq$ for the relation together with the semantic postulate that for all sentences φ and ψ:

$\varphi \preceq \psi$ is true iff for all θ, if $\psi \prec \theta$ is true, then $\varphi \prec \theta$ is true.

Thus, on the operator approach, we need to introduce additional syntax and additional semantics to deal with weak partial ground, while on the predicate approach we can use standard first-order definitions in the object language.

2.2 **Truth and Partial Ground**: A major benefit of the predicate approach is that it allows us to study the connections between partial ground and truth in a natural setting. It should be clear that partial ground is conceptually related to truth—partial ground is a relation on the truths after all. In axiomatic theories of truth, the concept is standardly formalized by means of a unary predicate of sentences [58]. By formalizing partial ground analogously using a relational predicate, we create a ground-theoretic framework in which we can fruitfully study the connections between truth and partial ground. For example, we will show in this paper that if we formulate the usually accepted principles for partial ground using a ground predicate, the resulting theory turns out to be a conservative extension of the well-known theory of positive truth [58, p. 116–22]. In other words, the predicate approach allows us to make the truth-theoretic commitments of theories of ground explicit. In the second part of this paper, we shall investigate the connections between partial ground and truth further. There we shall show, for example, that we can formulate a typed solution to Fine's *puzzle of ground* [43] in our axiomatic framework.

2.3 **Semantics of Partial Ground**: Finally, the ground predicate allows us to use classic model-theoretic methods to study the semantics of partial ground. In the literature on ground, we usually distinguish between *conceptualist* and *factualist* notions of ground [24, p. 256–59, 28, p. 14f]. On a conceptualist notion, ground is a relation on fine-grained, conceptually individuated *truths*. For example, on a conceptualist notion, we would typically say that if φ is a true sentence, then the truth of $\varphi \vee \varphi$ holds in virtue of the truth of φ, but not the other way around. On a factualist conception, in contrast, ground is a relation on coarse-grained, worldly individuated *facts*. On this conception, we would typically deny that if φ is a true sentence, the fact that $\varphi \vee \varphi$ holds in virtue of the fact that φ, since the two facts are the same—albeit expressed differently. The notion of partial ground that we are interested in here is a conceptualist notion of ground.

It is currently an open problem to provide a formal semantics for a conceptualist notion of ground.[11] On the operator approach, it is difficult to define such a semantics, since we have to start "from scratch," as it were: we have to find the right kind of structure to interpret conceptualist ground and provide primitive semantic clauses for the ground operator. On the predicate approach, in contrast, if we can develop a consistent first-order axiomatization of conceptualist ground, we can infer the existence of a (first-order) model by the completeness theorem for first-order logic. Once we know that such a model exists, we can study it using methods of classic model theory. This should then help us determine the right kind of structure and the correct semantic clauses to interpret conceptualist ground operators, as well.

In the rest of the paper, we will develop an axiomatization of partial ground over the truths of arithmetic, which fulfills the promises from the previous list of benefits. But before we begin, we shall briefly address an argument that is sometimes brought forward against the predicate approach: Correia [24, p. 254] and Fine [42, p. 46–47] argue that we should prefer the operator approach for reasons of *ontological neutrality*. They argue that since on the predicate approach we have terms denoting the relata of ground, by Quine's *criterion of ontological commitment*, the approach commits us to the existence of the relata of ground. Moreover, they argue that since on the predicate approach we are committed to the existence of the relata of ground, we need a background theory for them. On the operator approach, in contrast, they argue we don't have any of that: we only need to have the (true) sentences that the ground operator acts upon. This argument is particularly forceful on a factualist conception of partial ground, where we take the relata of ground to be *facts*. As Correia then puts it: "it should be possible to make claims of grounding and fail to believe in facts" [24, p. 254].

In this paper, however, we work on a conceptualist notion of partial ground, where we take the relata of ground to be *truths*.[12] Moreover, these truths are truths *of sentences*. Correspondingly, we formalize partial ground using a relational predicate of (true) sentences. Thus, by Quine's criterion of ontological commitment, we are only committed to the existence of (true)

[11]There are semantics for factualist notions of ground in the literature. The most commonly discussed semantics for the ground operator is given by Fine [42, p. 71–74] and Fine [44, p. 7–10] in terms of truthmakers. A related algebraic semantics is given by Correia [24, p. 274–76]. But as Fine [42, Fn 22, p. 74] himself notes, these semantics are not sound for a conceptualist notion of ground.

[12]For the distinction between truths and facts, see [41]. Note that according to Fine truths are not (true) sentences, rather they are derived entities that get their identity criteria from (true) sentences: truths according to Fine are a kind of linguistically individuated facts. In this paper, we don't presuppose a specific metaphysical understanding of truths: they can be anything from true sentences to metaphysically robust fact-like entities in their own right.

sentences. Our background theory is correspondingly simply a standard theory of syntax. These ontological commitments are metaphysically innocuous: ontologically speaking, sentences are relatively harmless entities. Moreover, on the operator approach, we need to assume the existence of sentences to formulate our theory in the first place. Thus, even though the operator approach is, strictly speaking, ontologically more parsimonious, the ontological commitments of our version of the predicate approach are fairly harmless.

2.4 Technical Preliminaries

To develop our axiomatization of partial ground over the truths of arithmetic, we need a background theory of arithmetic, which tells us what we need to know about arithmetic, and a background theory of syntax, which allows us to talk about the (true) sentences of arithmetic. It is well-known that PA can double as a theory of arithmetic and as a theory of syntax. This can be achieved using the technique of Gödel-numbering. In this section, we will recount the basics of this technique and fix notation.[13]

Let $\mathcal{L}$ be the language of PA. We assume that $\mathcal{L}$ has the standard arithmetic vocabulary: an individual constant 0 intended to denote the natural number *zero*, a unary function symbol S intended to express the *successor function* on the natural numbers, and binary function symbols $+$ and $\times$ intended to denote *addition* and *multiplication* on the natural numbers respectively. For every natural number n, we standardly define the *numeral* $\overline{n}$ as the n-fold application of S to the constant 0. The numeral $\overline{n}$ is, of course, intended to denote the number n. Note that '$\overline{n}$' is merely a meta-linguistic abbreviation of the official object-linguistic term '$S \dots Sn$'. The *language of truth* $\mathcal{L}_{Tr}$ is the result of extending $\mathcal{L}$ with the unary truth predicate Tr, the *language of predicational ground* $\mathcal{L}_{Tr}^{\lhd}$ is the result of extending $\mathcal{L}_{Tr}$ with the binary ground predicate $\lhd$, and the *language of (simple) operational ground* $\mathcal{L}_{\prec}$ is the result of extending $\mathcal{L}$ with the applications of the binary ground operator $\prec$ over $\mathcal{L}$: $\mathcal{L}_{\prec} := \mathcal{L} \cup \{\varphi \prec \psi \mid \varphi, \psi \in \mathcal{L}\}$.[14] In the following, we will mainly work in within $\mathcal{L}_{Tr}^{\lhd}$.

We use the technique of Gödel-numbering to obtain names for every expression. In particular, we use a *coding function* $\#$ to injectively map every expression σ to a natural number $\#\sigma$—the *Gödel number* of the expression. If σ is an expression, then we also write $\ulcorner \sigma \urcorner$ for the numeral intended to

[13]We assume that the reader is already familiar with the basics of first-order logic and has at least a rough understanding of how Gödel-numbering works. For the details, we refer the reader to [16].

[14]Note that we're explicitly excluding iterations of the operator $\prec$ here. See also our discussion on p. 52 below.

denote $\#\sigma$. This will be our *name* for σ. For the most part, we simply assume that we have some coding function for the language $\mathcal{L}$, but later we will discuss theories that require coding functions for $\mathcal{L}_{Tr}$ and even $\mathcal{L}_{Tr}^{\lhd}$.

The theory PA of PA consists of the standard axioms for zero, the successor function, addition, and multiplication, plus all the instances of the *induction scheme*

$$\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(Sx)) \to \forall x \varphi(x)$$

over formulas $\varphi(x)$ in the language $\mathcal{L}$. We denote derivability in PA by $\vdash_{PA}$ (and analogously for other systems discussed in the paper). However, if what we mean is clear from the context, we omit the subscript.

It is well-known that PA can represent any recursive function, in the sense that if f is a recursive function then there is a formula $\varphi(x, y)$ such that for all natural numbers n, m:

$$f(n) = m \text{ iff } \vdash_{PA} \forall x(\varphi(\overline{n}, x) \leftrightarrow x = \overline{m}).$$

Many syntactic functions on the codes of expressions are recursive and thus representable. For example, the function that maps the code $\#\varphi$ of a formula φ to the code $\#\neg\varphi$ of its negation is recursive. It is convenient to assume that $\mathcal{L}$ has function symbols for a finite number of those functions. Notation-wise, if f is a recursive function, then we use $\underset{\cdot}{f}$ as our function symbol for it. In particular, we assume that we have function symbols $\underset{\cdot}{\neg}$, $\underset{\cdot}{\vee}$, $\underset{\cdot}{\wedge}$, $\underset{\cdot}{\exists}$, $\underset{\cdot}{\forall}$, and $\underset{\cdot}{=}$ for the corresponding syntactic operations on the codes of expressions. If we work in the context of a coding for $\mathcal{L}_{Tr}$, we additionally assume a function symbol $\underset{\cdot}{Tr}$ for the function that maps the code $\#t$ of a term t to the code $\#Tr(t)$ of the atomic formula $Tr(t) \in \mathcal{L}_{Tr}$. And if we work in the context of a coding for $\mathcal{L}_{Tr}^{\lhd}$, we assume a function symbol $\underset{\cdot}{\lhd}$ for the function that maps the codes $\#s$ and $\#t$ of two terms to the code $\#(s \lhd t)$ of the atomic formula $s \lhd t$. We can then conservatively extend our axioms with the defining equations for those functions such that for all formulas φ and ψ, for all variables v, and for all terms t:

$$\vdash_{PA} \ulcorner s \urcorner \underset{\cdot}{=} \ulcorner t \urcorner = \ulcorner s = t \urcorner \qquad \vdash_{PA} \underset{\cdot}{\neg} \ulcorner \varphi \urcorner = \ulcorner \neg\varphi \urcorner$$

$$\vdash_{PA} \ulcorner \varphi \urcorner \underset{\cdot}{\wedge} \ulcorner \psi \urcorner = \ulcorner \varphi \wedge \psi \urcorner \qquad \vdash_{PA} \ulcorner \varphi \urcorner \underset{\cdot}{\vee} \ulcorner \psi \urcorner = \ulcorner \varphi \vee \psi \urcorner$$

$$\vdash_{PA} \underset{\cdot}{\exists}(\ulcorner v \urcorner, \ulcorner \varphi \urcorner) = \ulcorner \exists v \varphi \urcorner \qquad \vdash_{PA} \underset{\cdot}{\forall}(\ulcorner v \urcorner, \ulcorner \varphi \urcorner) = \ulcorner \forall v \varphi \urcorner$$

When we work in the context of coding functions for $\mathcal{L}_{Tr}$ and $\mathcal{L}_{Tr}^{\lhd}$, we furthermore get of all terms s and t that:

$$\vdash_{PA} \underset{\cdot}{Tr}(\ulcorner t \urcorner) = \ulcorner Tr(t) \urcorner \quad \vdash_{PA} \ulcorner s \urcorner \underset{\cdot}{\lhd} \ulcorner t \urcorner = \ulcorner s \lhd t \urcorner$$

Note that, in particular, we get that $\vdash_{PA} \underset{\cdot}{Tr}(\ulcorner \ulcorner \varphi \urcorner \urcorner) = \ulcorner Tr(\ulcorner \varphi \urcorner) \urcorner$, for every sentence φ. The ternary substitution function sub such that for all formulas φ, terms t, and variables v $sub(\#\varphi, \#t, \#v) = \#\varphi(t/v)$ provided that t is free for v in φ, is recursive and thus representable. Officially, we represent this function by the function symbol $\underset{.}{sub}$ and add its defining equations to our axioms, but unofficially we often simply write $\ulcorner\varphi\urcorner(\ulcorner t\urcorner, \ulcorner v\urcorner)$ instead of $\underset{.}{sub}(\ulcorner\varphi\urcorner, \ulcorner t\urcorner, \ulcorner v\urcorner)$ and if there is only one free variable in φ, we often simply write $\ulcorner\varphi\urcorner(\ulcorner t\urcorner)$. The function that maps a natural number n to the code $\#\overline{n}$ of its numeral $\overline{n}$ is also recursive and we will use the function symbol $\dot{}$ for this in our language. Note that, in particular, we get for all sentences φ that $\vdash_{PA} \ulcorner\dot{\varphi}\urcorner = \ulcorner\ulcorner\varphi\urcorner\urcorner$. We write $\ulcorner\varphi(\dot{x})\urcorner$ as an abbreviation for $\underset{.}{sub}(\ulcorner\varphi\urcorner, \dot{x})$. This allows us to quantify over free variables in the context of names. The *valuation function val* that applied to (the code of) a closed term yields its denotation is also recursive and thus representable. Officially, however, we cannot have a function symbol representing the valuation function, since otherwise we run the risk of inconsistency [58, p. 32]. We will nevertheless write $s^\circ = t$ to say that the denotation of s is t, as if $^\circ$ was a function symbol representing the valuation function. Officially, this is merely an abbreviation for the corresponding complex defining formula.

PA can also (strongly) represent every recursive set, in the sense that if S is a recursive set then there is a formula $\varphi(x)$ such that for all natural numbers n:

$$n \in S \text{ iff } \vdash_{PA} \varphi(\overline{n}) \quad \text{and} \quad n \notin S \text{ iff } \vdash_{PA} \neg\varphi(\overline{n})$$

In the following, we'll write $Sent$ to abbreviate the formula that allows us to represent the recursive set of codes sentences in $\mathcal{L}$, $Sent_{Tr}$ for the formula that allows us to represent the codes of sentences in $\mathcal{L}_{Tr}$, and $Sent^{\lhd}_{Tr}$ for the formula that allows us to represent the codes of sentences in $\mathcal{L}^{\lhd}_{Tr}$. Similarly, Var and $ClTerm$ are abbreviations for the formulas that allow us to represent the sets of (codes of) variables and closed terms. As an abbreviation for $\forall x(Var(x) \rightarrow \varphi(x))$ we write $\forall v\varphi(v)$ and as an abbreviation for $\forall x(ClTerm(x) \rightarrow \varphi(x))$ we write $\forall t\varphi(t)$. We also sometimes use the notation $\forall tTr(\ulcorner\varphi(\underset{.}{t})\urcorner)$ for $\forall x(ClTerm(x) \rightarrow Tr(\underset{.}{sub}(\ulcorner\varphi\urcorner, x)))$. This allows us to quantify over terms in the context of names.

We assume that PA has the defining axioms for all of these function symbols and predicates as axioms. Furthermore, the theory PAT extends PA with the missing instances of the induction scheme over $\mathcal{L}_{Tr}$ and the theory PAG extends PAT with all the missing instances of the induction scheme over $\mathcal{L}^{\lhd}_{Tr}$.

Finally, we will exclusively work in the context of the *standard model* of PA. This model of $\mathcal{L}$ has the set $\mathbb{N}$ of the natural numbers as its domain and in it 0 actually denotes the number zero, S actually denotes the successor function, and $+$ and $\times$ actually denote addition and multiplication. In other words, we don't allow for non-standard interpretations of the arithmetic vocabulary. We denote this model by $\mathbf{N}$. A model for $\mathcal{L}_{Tr}$, then, has the form $(\mathbf{N}, S)$, where $\mathbf{N}$ is the standard model and $S \subseteq \mathbb{N}$ interprets the truth predicate Tr. A model for $\mathcal{L}_{Tr}^{\triangleleft}$ has the form $(\mathbf{N}, S, R)$, where $(\mathbf{N}, S)$ is a model of $\mathcal{L}_{Tr}$ and $R \subseteq \mathbb{N}^2$ interprets our ground predicate $\triangleleft$. Thus, on our notion of a model, the interpretation of the arithmetic vocabulary is fixed, but we are allowed to freely interpret the truth predicate and the ground predicate.[15] Note that we don't have a notion of a model of $\mathcal{L}_{\prec}$, since finding appropriate models for this language is an open problem.

2.5 Axiomatic Theories of Partial Ground

2.5.1 Axioms for Partial Ground

We begin from the standardly accepted principles for partial ground formulated on the operator approach. The most comprehensive conceptualist system for ground on the operator approach is the *pure and impure logic of ground* developed by Fine [42, p. 54–71]. However, Fine's system deals with various notions of ground and takes the *full* notion of ground as fundamental [42, p. 50]—it contains a system for partial ground only as a subsystem. Moreover, Fine's system is formulated in a *sequent-style*, which makes it difficult to deal with for our present purpose. For these reasons, we will take the system of Schnieder [124] as our starting point. Schnieder's system is not primarily intended as a system for partial ground: it is intended as a system for the non-causal uses of the binary explanatory connective 'because' from natural language [124, Fn 8, p. 446–47]. However, there are uses of 'because' that coincide with the present sense of partial ground: when we say that one truth holds either wholly or partially *because of* another truth, we can interpret this as saying that the one truth holds either wholly or partially *in virtue of* the other truth. The interpretation of 'because' is often given in the literature on ground and is sometimes even used as a paradigmatic natural language example for ground [42, p. 37–38].[16] Since Schnieder's system is supposed to account for *all* non-causal uses of 'because', it should also cover *this* non-causal use of because—in other words: we can interpret Schnieder's system as a system for partial ground.[17]

Schnieder formulates his system over pure first-order logic as his base-theory, but the system can easily be adapted to the present framework. If we take

[15] The notation and background theory we use in this paper is adapted from the standard notation and background theory used in axiomatic theories of truth. For the reader not familiar with these conventions, we recommend [58, p. 29–38].

[16] For a detailed discussion of the relation between 'because' and 'in virtue of', see [15, §4].

[17] In fact, we can show that the fragment of Fine's system that deals with partial ground coincides with Schnieder's system interpreted as a system for partial ground.

Schnieder's system and formulate it in the language $\mathcal{L}_\prec$ over PA as its base-theory, we arrive at the following system:

Definition 2.5.1. *The operational theory of (partial) ground OG consists of the axioms of PA, all the instances of the axiom scheme:*

$$\neg(\varphi \prec \varphi),$$

for sentences $\varphi \in \mathcal{L}$, plus the following rules of inference for partial ground for all formulas $\varphi, \psi, \theta \in \mathcal{L}$:

$$\frac{\varphi \prec \psi \quad \psi \prec \theta}{\varphi \prec \theta} \qquad \frac{\varphi \prec \psi}{\varphi} \qquad \frac{\varphi \prec \psi}{\psi} \qquad \frac{\varphi}{\varphi \prec \neg\neg\varphi}$$

$$\frac{\varphi}{\varphi \prec \varphi \vee \psi} \qquad \frac{\psi}{\psi \prec \varphi \vee \psi} \qquad \frac{\varphi \quad \psi}{\varphi \prec \varphi \wedge \psi} \qquad \frac{\varphi \quad \psi}{\psi \prec \varphi \wedge \psi}$$

$$\frac{\neg\varphi}{\neg\varphi \prec \neg(\varphi \wedge \psi)} \qquad \frac{\neg\psi}{\neg\psi \prec \neg(\varphi \wedge \psi)} \qquad \frac{\neg\varphi \quad \neg\psi}{\neg\varphi \prec \neg(\varphi \vee \psi)} \qquad \frac{\neg\varphi \quad \neg\psi}{\neg\psi \prec \neg(\varphi \vee \psi)}$$

$$\frac{\forall x\varphi(x)}{\varphi(t) \prec \forall x\varphi(x)} \qquad \frac{\varphi(t)}{\varphi(t) \prec \exists x\varphi(x)} \qquad \frac{\neg\varphi(t)}{\neg\varphi(t) \prec \neg\forall x\varphi(x)} \qquad \frac{\forall x\neg\varphi(x)}{\neg\varphi(t) \prec \neg\exists x\varphi(x)}$$

Note well that the theory OG is formulated in the language $\mathcal{L}_\prec$, which explicitly doesn't allow for iterations of $\prec$. This is in line with the standard restriction in the literature to un-iterated or *simple* instances of ground. Iterated ground raises specific technical and philosophical issues, which fall outside the scope of this article.[18]

Schnieder [124, p. 452–53] shows the proof-theoretic conservativity of the propositional fragment of his system over pure propositional logic. This proof is easily extended to show the conservativity of his quantified system, which we used as our starting point, over pure first-order logic.[19] However, since we take PA as our background theory, we give a slightly different proof of the analogous result for the present context:

Proposition 2.5.2 (Schnieder)**.** *The system OG is a proof-theoretically conservative extension of PA.*

Proof. The *complexity function c*, which maps the code $\#\varphi$ of a formula φ to the code $\#|\varphi|$ of its logical complexity $|\varphi|$, is recursive and thus representable in PA. Let $\underset{\cdot}{c}$ represent this function. Furthermore, let $\underset{\cdot}{<}$ represent

[18]For a discussion of these issues, see, e.g., [11, 32, 85]. There are particular issues to do with iterated ground that arise in the context of the predicate approach taken in this article, which will be explicitly discussed in the second part of the article.

[19]Schnieder does not carry out the details himself. The proof is left to the interested reader.

the recursive *strictly-less-than relation* $<$ on the natural numbers. We define the translation function $\tau : \mathcal{L}_{\prec} \to \mathcal{L}$ recursively by saying that:

(i) $\tau(\varphi) = \varphi$, for φ an atomic formula;

(ii) $\tau(\neg\varphi) = \neg\tau(\varphi)$;

(iii) $\tau(\varphi \circ \psi) = \tau(\varphi) \circ \tau(\psi)$, for $\circ = \wedge, \vee$;

(iv) $\tau(Qx\varphi) = Qx(\tau(\varphi))$, for $Q = \forall, \exists$; and

(v) $\tau(\varphi \prec \psi) = (\varphi \wedge \psi \wedge \underset{\cdot}{c}(\ulcorner\varphi\urcorner) \underset{\cdot}{<} \underset{\cdot}{c}(\ulcorner\psi\urcorner))$.

Note that in clause (v), we need not translate φ and ψ, since they are, by assumption, already in $\mathcal{L}$.

It is now easily seen by induction on the complexity of formulas that (a) for all $\varphi \in \mathcal{L}$, $\tau(\varphi) = \varphi$. In words: τ is constant on the arithmetic formulas. Next, we show that (b) τ preserves theoremhood over the two systems OG and PA, in the sense that for all $\varphi \in \mathcal{L}_{\prec}$, if $\vdash_{OG} \varphi$, then $\vdash_{PA} \tau(\varphi)$. We show (b) by an induction on the length of derivations. Of course, we only need to consider the rules of OG that are not rules of PA. If $\varphi = \neg(\varphi' \prec \varphi')$ is an instance of the axiom scheme of OG, for $\varphi' \in \mathcal{L}$, then we get that $\tau(\varphi) =_{\text{(v),(a)}} \neg(\varphi' \wedge \varphi' \wedge \underset{\cdot}{c}(\varphi') \underset{\cdot}{<} \underset{\cdot}{c}(\varphi'))$. But $\vdash_{PA} \neg(\varphi' \wedge \varphi' \wedge \underset{\cdot}{c}(\varphi') \underset{\cdot}{<} \underset{\cdot}{c}(\varphi'))$, since $\vdash_{PA} \forall x \neg(x \underset{\cdot}{<} x)$ and thus in particular $\vdash_{PA} \neg(\underset{\cdot}{c}(\varphi') \underset{\cdot}{<} \underset{\cdot}{c}(\varphi'))$. So assume the induction hypothesis. For the induction step, we need to go through all the inference rules of OG case by case. Here we only discuss one case to illustrate the idea. Consider the case where the last step has been an application of the rule:

$$\frac{\varphi}{\varphi \prec \neg\neg\varphi} \ ,$$

where $\varphi \in \mathcal{L}$. First note that since $\varphi \in \mathcal{L}$, we get that $\tau(\varphi) = \varphi$ by (a) and furthermore that $(*)$ $\vdash_{PA} \varphi$ by the induction hypothesis. Now consider, $\tau(\varphi \prec \neg\neg\varphi) =_{(v),(a)} \varphi \wedge \neg\neg\varphi \wedge \underset{\cdot}{c}(\ulcorner\varphi\urcorner) \underset{\cdot}{<} \underset{\cdot}{c}(\underset{\cdot}{\neg}\underset{\cdot}{\neg}\ulcorner\varphi\urcorner)$. By $(*)$ we know that $\vdash_{PA} \varphi$ and thus $\vdash_{PA} \neg\neg\varphi$ by elementary logic. And we know that $\vdash_{PA} \underset{\cdot}{c}(\ulcorner\varphi\urcorner) \underset{\cdot}{<} \underset{\cdot}{c}(\underset{\cdot}{\neg}\underset{\cdot}{\neg}\ulcorner\varphi\urcorner)$, since $\vdash_{PA} \forall x(Sent(x) \to \underset{\cdot}{c}(x) \underset{\cdot}{<} \underset{\cdot}{c}(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x))$. The other cases are equally straightforward. Putting the two claims (a) and (b) together, the proposition follows. $\square$

$\square$

Note that the translation function used in the proof is not particularly faithful: we can derive a lot of intuitively false claims under the translation. For example, it is intuitively false that $0 = 0 \prec \neg\exists x(Sx = 0)$, since the (logical) truth of $0 = 0$ has nothing to do with the truth of (the axiom) $\neg\exists x(Sx = 0)$.

Nevertheless, we will get that $\vdash_{PA} \tau(0 = 0 \prec \neg\exists x(Sx = 0))$ with τ defined as in the proof, since $\tau(0 = 0 \prec \neg\exists x(Sx = 0))$ is equal to

$$0 = 0 \wedge \neg\exists x(Sx = 0) \wedge c\!\!\!{}_{\cdot}(\ulcorner 0 = 0 \urcorner) \mathrel{\underset{\cdot}{<}} c\!\!\!{}_{\cdot}(\ulcorner \neg\exists x(Sx = 0)\urcorner),$$

which is provable in PA. But this is not a counterexample to the claim in the proof, since the operational theory of ground does not prove $0 = 0 \prec \neg\exists x(Sx = 0)$ to begin with, and all that is required for our proof is that the translation preserves theoremhood. The result, then, immediately gives us the proof-theoretic consistency of the predicational theory of partial ground:

Corollary 2.5.3 (Schnieder)**.** *The system OG is proof-theoretically consistent.*

Note that the proof-theoretic consistency of the operational theory of ground does not entail that there are models for the theory, since we have not even defined the notion of a model for its language, much less have we shown that proof-theoretic consistency in this language implies the existence of models.

We obtain our axiomatization of partial ground over the truths of arithmetic by translating the axioms and rules of the operational theory into quantified axioms, which we formulate using the ground predicate. For this purpose, we assume that we have a Gödel-numbering for $\mathcal{L}$. Let us begin with the axiom scheme $\neg(\varphi \prec \varphi)$, for sentences $\varphi \in \mathcal{L}$, which expresses the irreflexivity of partial ground. We straight-forwardly translate this to the quantified axiom $\forall x \neg(x \lhd x)$. To translate the first rule:

$$\frac{\varphi \prec \psi \quad \psi \prec \theta}{\varphi \prec \theta}\ ,$$

which captures the transitivity of partial ground, we first transform the rule into the conditional $(\varphi \prec \psi) \wedge (\psi \prec \theta) \to (\varphi \prec \theta)$, and then translate this conditional into the quantified axiom $\forall x \forall y \forall z(x \lhd y \wedge y \lhd z \to x \lhd z)$. To translate the remaining rules, we need to use a "trick" in order to quantify over formulas that are affirmed outside the context of the ground operator. Take the rules:

$$\frac{\varphi \prec \psi}{\varphi} \quad \text{and} \quad \frac{\varphi \prec \psi}{\psi}\ ,$$

which express that partial ground is a relation on the truths. Again, we first translate the rules into the conditionals $(\varphi \prec \psi) \to \varphi$ and $(\varphi \prec \psi) \to \psi$. Then, in order to quantify over the formulas affirmed in the consequent, we use the truth predicate Tr. With some simplification, we get the following transformation:

$$\left.\begin{array}{c} \dfrac{\varphi \prec \psi}{\varphi} \\[2ex] \dfrac{\varphi \prec \psi}{\psi} \end{array}\right\} \quad \rightsquigarrow \quad \forall x \forall y (x \lhd y \rightarrow Tr(x) \wedge Tr(y))$$

Note that we use the truth predicate Tr here simply as a quantificational device: it allows us to generalize over the truths involved in partial ground outside the context of partial ground.

By applying the same strategy to the rule involving double negation, we get the following transformation:

$$\frac{\varphi}{\varphi \prec \neg\neg\varphi} \quad \rightsquigarrow \quad \forall x (Tr(x) \rightarrow x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$$

But now we face a problem: The operational theory of ground can not only prove that if φ is a true sentence, then the truth of $\neg\neg\varphi$ is grounded in the truth of φ, but also that if $\neg\neg\varphi$ is a true sentence, then the truth of $\neg\neg\varphi$ is grounded in the truth of φ. Formally, we get both $\vdash_{OG} \varphi \rightarrow (\varphi \prec \neg\neg\varphi)$ and $\vdash_{OG} \neg\neg\varphi \rightarrow (\varphi \prec \neg\neg\varphi)$. But the corresponding quantified claim $\forall x(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \rightarrow x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$ is *not* derivable from our axioms so far. In response to this, we might be tempted to simply add the T-scheme:

$$Tr(\ulcorner \varphi \urcorner) \leftrightarrow \varphi,$$

for all sentences $\varphi \in \mathcal{L}$, to our theory. This would allow us to derive $Tr(\ulcorner \neg\neg\varphi \urcorner) \rightarrow \ulcorner \varphi \urcorner \lhd \ulcorner \neg\neg\varphi \urcorner$ for every formula φ. But this is not enough. We wish to derive the full *quantified* claim $\forall x(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \rightarrow x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$ in our theory and merely using the T-scheme this is impossible. Therefore, we will add the quantified claim as an axiom to our system.

Thus, corresponding to every rule of the operational theory of ground, we have two axioms: an *upward directed* axiom, like $\forall x(Tr(x) \rightarrow x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$, and a *downward directed* axiom, like $\forall x(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \rightarrow x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$. In the cases of the other rules, we can moreover make some simplifications. To illustrate, consider the case of the rules involving conjunction. We get the following transformations:

Upward:

$$\left.\begin{array}{c} \dfrac{\varphi \quad \psi}{\varphi \prec \varphi \wedge \psi} \\[2ex] \dfrac{\varphi \quad \psi}{\psi \prec \varphi \wedge \psi} \end{array}\right\} \quad \rightsquigarrow \quad \forall x \forall y (Tr(x) \wedge Tr(y) \rightarrow x \lhd x \underset{\cdot}{\wedge} y \wedge y \lhd x \underset{\cdot}{\wedge} y)$$

Downward:

$$\varphi \wedge \psi \rightarrow (\varphi \prec \varphi \wedge \psi) \wedge (\psi \prec \varphi \wedge \psi) \rightsquigarrow \forall x \forall y (Tr(x \dot{\wedge} y) \rightarrow x \lhd x \dot{\wedge} y \wedge y \lhd x \dot{\wedge} y)$$

Intuitively, the upward directed axioms say what truths a given truth grounds, while the downward directed axioms say what truths ground a given truth. And intuitively, both kinds of principles are required: we want to say both what grounds a truth and in what it is grounded.[20]

Finally, we need to add some axioms to get our hierarchy "off the ground," as it were. So far, we have no axiom that allows us to introduce the truth predicate. Thus, we are not able to prove the antecedent of any of our quantified axioms. To fix this, we propose to add some basic truth axioms that apply the truth predicate to the atomic formulas of PA—they allow us to introduce the truth predicate in the case of true equations. For this purpose, we use the standard idea from axiomatic theories of truth:

$$\forall s \forall t (Tr(s \dot{=} t) \leftrightarrow s^\circ = t^\circ)$$
$$\forall s \forall t (Tr(s \dot{\neq} t) \leftrightarrow s^\circ \neq t^\circ)^{21}$$

In words, an equation is true iff the terms flanking the equality symbol have the same denotation. These axioms get our axiomatization "off the ground," in the sense that we can prove the truth of true equations and then use the other axioms to track partial ground through the complexity of the truths. Finally, since we wish to talk about the truths of arithmetic, we want to ensure that the truth predicate only applies to sentences of $\mathcal{L}$. We achieve this by postulating that:

$$\forall x (Tr(x) \rightarrow Sent(x)).$$

Note that from this axiom together with the axiom $\forall x \forall y (x \lhd y \rightarrow Tr(x) \wedge Tr(y))$ it follows that also the ground predicate only applies to sentences of $\mathcal{L}$:

$$\forall x \forall y (x \lhd y \rightarrow Sent(x) \wedge Sent(y)).$$

In other words, out theory is a simply typed theory of partial ground.

We arrive at the following axiomatization:

Definition 2.5.4. *The predicational theory of (partial) ground PG consists of the axioms of PAG plus the following axioms:*

[20] A nice feature of our theory is that it proves all the instances of the T-scheme via the upward and downward directed axioms. We will show this in §2.5.2. The point is that the upward and downward directed axioms are intuitively motivated and *on top of that* they give us back the T-scheme.

[21] Here $s \neq t$, for terms s and t, is an abbreviation of $\neg(s = t)$. Correspondingly, the notation $s \dot{\neq} t$ is an abbreviation for the complex function term $\dot{\neg}(s \dot{=} t)$, for terms s and t.

Basic Ground Axioms:

G_1 $\forall x \neg(x \lhd x)$

G_2 $\forall x \forall y \forall z(x \lhd y \wedge y \lhd z \rightarrow x \lhd z)$

G_3 $\forall x \forall y(x \lhd y \rightarrow Tr(x) \wedge Tr(y))$

Basic Truth Axioms:

T_1 $\forall s \forall t(Tr(s \underset{\cdot}{=} t) \leftrightarrow s^\circ = t^\circ)$

T_2 $\forall s \forall t(Tr(s \underset{\cdot}{\neq} t) \leftrightarrow s^\circ \neq t^\circ)$

T_3 $\forall x(Tr(x) \rightarrow Sent(x))$

Upward Directed Axioms:

U_1 $\forall x(Tr(x) \rightarrow x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg} x)$

U_2 $\forall x \forall y((Tr(x) \rightarrow x \lhd x \underset{\cdot}{\vee} y) \wedge (Tr(y) \rightarrow y \lhd x \underset{\cdot}{\vee} y))$

U_3 $\forall x \forall y(Tr(x) \wedge Tr(y) \rightarrow (x \lhd x \underset{\cdot}{\wedge} y) \wedge (y \lhd x \underset{\cdot}{\wedge} y))$

U_4 $\forall x \forall y(Tr(\underset{\cdot}{\neg} x) \wedge Tr(\underset{\cdot}{\neg} y) \rightarrow (\underset{\cdot}{\neg} x \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)) \wedge (\underset{\cdot}{\neg} y \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)))$

U_5 $\forall x \forall y((Tr(\underset{\cdot}{\neg} x) \rightarrow \underset{\cdot}{\neg} x \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\wedge} y)) \wedge (Tr(\underset{\cdot}{\neg} y) \rightarrow \underset{\cdot}{\neg} y \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\wedge} y)))$

U_6 $\forall x \forall t \forall v(Tr(x(t/v)) \rightarrow x(t/v) \lhd \underset{\cdot}{\exists} vx)$

U_7 $\forall x \forall v(\forall t Tr(\underset{\cdot}{\neg} x(t/v)) \rightarrow \forall t(\underset{\cdot}{\neg} x(t/v) \lhd \underset{\cdot}{\neg}\underset{\cdot}{\exists} vx))$

U_8 $\forall x \forall v(\forall t(Tr(x(t/v)) \rightarrow \forall t(x(t/v) \lhd \underset{\cdot}{\forall} vx))$

U_9 $\forall x \forall t \forall v(Tr(\underset{\cdot}{\neg} x(t/v)) \rightarrow \underset{\cdot}{\neg} x(t/v) \lhd \underset{\cdot}{\neg}\underset{\cdot}{\forall} vx))$

Downward Directed Axioms:

D_1 $\forall x(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg} x) \rightarrow x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg} x)$

D_2 $\forall x \forall y(Tr(x \underset{\cdot}{\vee} y) \rightarrow (Tr(x) \rightarrow x \lhd x \underset{\cdot}{\vee} y) \wedge (Tr(y) \rightarrow y \lhd x \underset{\cdot}{\vee} y))$

D_3 $\forall x \forall y(Tr(x \underset{\cdot}{\wedge} y) \rightarrow (x \lhd x \underset{\cdot}{\wedge} y) \wedge (y \lhd x \underset{\cdot}{\wedge} y))$

D_4 $\forall x \forall y(Tr(\underset{\cdot}{\neg}(x \underset{\cdot}{\wedge} y)) \rightarrow (Tr(\underset{\cdot}{\neg} x) \rightarrow \underset{\cdot}{\neg} x \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)) \wedge (Tr(\underset{\cdot}{\neg} y) \rightarrow \underset{\cdot}{\neg} y \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)))$

D_5 $\forall x \forall y(Tr(\underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)) \rightarrow (\underset{\cdot}{\neg} x \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)) \wedge (\underset{\cdot}{\neg} y \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)))$

D_6 $\forall x(Tr(\underset{\cdot}{\exists} vx(v)) \rightarrow \exists t(x(t/v) \lhd \underset{\cdot}{\exists} vx))$

D_7 $\forall x \forall v(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\exists} vx) \rightarrow \forall t(\underset{\cdot}{\neg} x(t/v) \lhd \underset{\cdot}{\neg}\underset{\cdot}{\exists} vx))$

D_8 $\forall x \forall v(Tr(\underset{\cdot}{\forall} vx \rightarrow \forall t(x(t/v) \lhd \underset{\cdot}{\forall} vx))$

D_9 $\forall x \forall v(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\forall} vx) \rightarrow \exists t(\underset{\cdot}{\neg} x(t/v) \lhd \underset{\cdot}{\neg}\underset{\cdot}{\forall} vx))$

2.5.2 Conservativity and Truth-theoretic Commitments

We will now determine the truth-theoretic commitments and the proof-theoretic strength of our predicational theory of ground.

As we have claimed in §2.3.2, the predicational theory of ground proves the well-known theory of positive truth [58, p.116–22]:

Definition 2.5.5 ('Positive Truth'). *The theory PT of positive truth is formulated in $\mathcal{L}_{Tr}$ and consists of the axioms of PAT and the three base truth axioms T_1, T_2, and T_3, plus the following axioms:*

P_1 $\forall x(Tr(x) \leftrightarrow Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x))$

P_2 $\forall x\forall y(Tr(x\underset{\cdot}{\land}y) \leftrightarrow Tr(x) \land Tr(y))$

P_3 $\forall x\forall y(Tr(\underset{\cdot}{\neg}(x\underset{\cdot}{\land}y)) \leftrightarrow Tr(\underset{\cdot}{\neg}x \lor T\underset{\cdot}{\neg}y))$

P_4 $\forall x\forall y(Tr(x\underset{\cdot}{\lor}y) \leftrightarrow Tr(x) \lor Tr(y))$

P_5 $\forall x\forall y(Tr(\underset{\cdot}{\neg}(x\underset{\cdot}{\lor}y)) \leftrightarrow Tr(\underset{\cdot}{\neg}x) \land Tr(\underset{\cdot}{\neg}y))$

P_6 $\forall x\forall v(Tr(\underset{\cdot}{\forall}vx) \leftrightarrow \forall t Tr(x(t/v)))$

P_7 $\forall x\forall v(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\forall}vx) \leftrightarrow \exists t Tr(\underset{\cdot}{\neg}x(t/v)))$

P_8 $\forall x\forall v(Tr(\underset{\cdot}{\exists}vx) \leftrightarrow \exists t Tr(x(t/v)))$

P_9 $\forall x\forall v(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\exists}vx) \leftrightarrow \forall t Tr(\underset{\cdot}{\neg}x(t/v)))$

Proposition 2.5.6. *$PG \vdash PT$.*

Proof. It suffices to derive P_{1-9}. The derivation proceeds in every case by putting the upward directed and the downward directed axioms of PG together while using the axiom G_3. Here we only sketch the derivation of axiom P_1:

1. $\forall x(Tr(x) \to x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$ (U_1)
2. $\forall x\forall y(x \lhd y \to Tr(x) \land Tr(y))$ (G_3)
3. $\forall x(Tr(x) \to Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x))$ (1., 2., $\to$-Elim, and $\land$-Elim)
4. $\forall x(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \to x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$ (D_1)
5. $\forall x(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \to Tr(x))$ (2., 4., $\to$-Elim, and $\land$-Elim)
6. $\forall x(Tr(x) \leftrightarrow Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x))$ (3., 5., $\leftrightarrow$-Intro)

The other axioms can be derived analogously. □ □

This result connects the debate about axiomatic theories of truth with the debate about partial ground. In particular, the result shows that once we move to a shared framework for theories of truth and theories of ground, we only need to accept a theory of partial ground to get a proper theory of truth. Admittedly, this theory has to be formulated using the truth predicate, but this use is intuitively justified, since we have used the truth predicate simply as a quantificational device in formalizing the principles for partial ground.[22]

[22] It is well-known that the theory of positive truth has the same theorems as the theory of compositional truth: **Definition**: *The theory CT of compositional truth has the axioms of PAT, the two basic truth axioms T_1 and T_3 plus the following axioms:*

Proposition 2.5.6 has several interesting consequences. First, it entails as a simple corollary that our predicational theory of ground proves all the instances of the T-scheme:

Corollary 2.5.7. *PG proves the uniform T-scheme for all sentences $\varphi \in \mathcal{L}$:*

$$\vdash_{PG} \forall t_1, \ldots, \forall t_n(Tr(\varphi(t_{\dot{1}}, \ldots, t_{\dot{n}})) \leftrightarrow \varphi(t_1^{\circ}, \ldots, t_n^{\circ}))$$

Proof. It is well-known that PT proves all the instances of the uniform T-scheme over $\mathcal{L}$. The proof of this proceeds by a simple induction on the positive complexity of formulas. Our claim follows by Proposition 2.5.6. □

□

Second, we can use Proposition 2.5.6 and its Corollary 2.5.7 to show interesting facts about our predicational theory of ground, such as the fact that ground is *hyperintensional* according to our theory. Remember that a context is hyperintensional iff in the context the substitution of logical equivalents need not preserve truth-value (Cresswell [29]). Now, the set of (codes of) *logical truths of* $\mathcal{L}_{Tr}^{\lhd}$ is recursively enumerable, and thus weakly representable in the PA, i.e. there is a formula $\varphi(x)$ such that:

$$\vdash_{PA} \varphi(\overline{n}) \text{ iff } n \text{ is the code of a logical truth.}$$

Let's abbreviate this formula $\varphi(x)$ by $\mathsf{Val}(x)$. In particular, we'll get for all formulas $\varphi \in \mathcal{L}_{Tr}^{\lhd}$:

$$\vdash_{PA} \mathsf{Val}(\ulcorner\varphi\urcorner) \text{ iff } \varphi \text{ is a logical truth of } \mathcal{L}_{Tr}^{\lhd}.$$

With this preliminaries in place, we can show the following result:

Lemma 2.5.8. *PG proves that partial ground is hyperintensional in the following sense:*

(i) $\vdash_{PG} \neg\forall x\forall y(\mathsf{Val}(x \dot{\leftrightarrow} y) \to \forall z(z \lhd x \leftrightarrow z \lhd y))$

(ii) $\vdash_{PG} \neg\forall x\forall y(\mathsf{Val}(x \dot{\leftrightarrow} y) \to \forall z(x \lhd z \leftrightarrow y \lhd z))$

C_1 $\forall x(Tr(\dot{\neg}x) \leftrightarrow \neg Tr(x))$

C_2 $\forall x\forall y(Tr(x \dot{\vee} y) \leftrightarrow Tr(x) \vee Tr(y))$

C_3 $\forall x\forall y(Tr(x \dot{\wedge} y) \leftrightarrow Tr(x) \wedge Tr(y))$

C_4 $\forall v\forall x(Tr(\dot{\exists}vx) \leftrightarrow \exists t Tr(x(t/v)))$

C_5 $\forall v\forall x(Tr(\dot{\forall}vx) \leftrightarrow \forall t Tr(x(t/v)))$

For a proof of the equivalence of PT and CT, see [58, p. 120]. Note that the result depends on the fact that we start from the theory PAT in defining both PT and CT. The theories defined by the same axioms over PA as their base theory are *not* equivalent [see 58, p. 120]. In the following, we will often use well-known results about CT and apply them immediately to PT.

Proof. Note, for example, that $\vdash_{PG}$ $\mathsf{Val}(\ulcorner 0 = 0 \urcorner \underset{\cdot}{\leftrightarrow} \underset{\cdot}{\neg}\underset{\cdot}{\neg}\ulcorner 0 = 0 \urcorner)$ and $\vdash_{PG}$ $\ulcorner 0 = 0 \urcorner \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}\ulcorner 0 = 0 \urcorner$, but $\vdash_{PG} \neg(\ulcorner 0 = 0 \urcorner \lhd \ulcorner 0 = 0 \urcorner)$ establishing (i) and $\vdash_{PG} \neg(\underset{\cdot}{\neg}\underset{\cdot}{\neg}\ulcorner 0 = 0 \urcorner \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}\ulcorner 0 = 0 \urcorner)$ establishing (ii). $\square$ $\square$

Note that the proof of Lemma 2.5.8 makes use of facts that we get from Proposition 2.5.6 and Corollary 2.5.7 in several places. Without using these facts, the lemma would be difficult to prove.

Moreover, we can show a kind of adequacy result for our axiomatization with respect to the operational theory of ground:

Proposition 2.5.9. *For all sentences $\varphi, \psi \in \mathcal{L}$, if $\vdash_{OG} \varphi \prec \psi$, then $\vdash_{PG}$ $\ulcorner \varphi \urcorner \lhd \ulcorner \psi \urcorner$.*

Proof. By an induction on the length of derivations in OG. The only interesting step is when the last inference in a derivation was an application of an inference rule for the partial ground operator. Consider the step where the last inference was of the form:

$$\frac{\varphi}{\varphi \prec \neg\neg\varphi}$$

This means that $\vdash_{OG} \varphi$. Since $\varphi \in \mathcal{L}$, we get by Proposition 2.5.2 that $\vdash_{PA} \varphi$ and thus $\vdash_{PG} \varphi$. Then by applying the T-scheme we get $\vdash_{PG} Tr(\ulcorner \varphi \urcorner)$ and from this, using the axiom U_1, $\vdash_{PG} \ulcorner \varphi \urcorner \lhd \ulcorner \neg\neg\varphi \urcorner$. The cases for the other rules are analogous. Note that we only need the induction hypothesis in the case of the rule:

$$\frac{\varphi \prec \psi \quad \psi \prec \theta}{\varphi \prec \theta}\ ,$$

which is the only rule that has formulas with the ground operator in its antecedent. We get the result immediately by applying the induction hypothesis to the antecedents and using the axioms G_2, which captures the transitivity of partial ground. $\square$ $\square$

The other direction of Proposition 2.5.9, however, does *not* hold. This follows from the fact that the theory of positive truth and thus the predicational theory of ground is not conservative over PA. For example, positive truth proves the consistency of PA, in the sense that $\vdash_{PT} \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)$.[23] Thus, by Proposition 2.5.6, $\vdash_{PG} \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)$. But by Gödel's second incompleteness theorem, we know that $(Incomp)$ $\nvdash_{PA} \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)$.

[23] The unary predicate Bew_{PA} strongly represents the set of codes of sentences provable in PA. Here we simply take it to be an abbreviation of the (long) defining formula for Bew.

Now, since $\vdash_{PG} \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)$, we know using the T-scheme that $\vdash_{PG} Tr(\ulcorner \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)\urcorner)$. Using the axiom U_5, we can infer from this that:

$$\vdash_{PG} \ulcorner \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)\urcorner \lhd \ulcorner \neg (Bew_{PA}(\ulcorner 0 = 1 \urcorner) \wedge Bew_{PA}(\ulcorner 0 = 1 \urcorner))\urcorner.$$

But we cannot have that:

$$\vdash_{OG} \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner) \prec \neg (Bew_{PA}(\ulcorner 0 = 1 \urcorner) \wedge Bew_{PA}(\ulcorner 0 = 1 \urcorner)),$$

since then we would get the following derivation in the operational theory of ground:

$$\dfrac{\begin{array}{c}\vdots\\ \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner) \prec \neg (Bew_{PA}(\ulcorner 0 = 1 \urcorner) \wedge Bew_{PA}(\ulcorner 0 = 1 \urcorner))\end{array}}{\neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)} \; ,$$

by means of the OG rule

$$\frac{\varphi \prec \psi}{\varphi} \; ,$$

applied to the supposed derivation of $\neg Bew_{PA}(\ulcorner 0 = 1 \urcorner) \prec \neg (Bew_{PA}(\ulcorner 0 = 1 \urcorner) \wedge Bew_{PA}(\ulcorner 0 = 1 \urcorner))$. This would then mean that $\vdash_{OG} \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)$ and thus by Proposition 2.5.2, since $\neg Bew_{PA}(\ulcorner 0 = 1 \urcorner) \in \mathcal{L}$, that $\vdash_{PA} \neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)$—which is in contradiction to $(Incomp)$. Moreover, intuitively speaking, the sentence $\neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)$ is true—we know, for example by Gentzen's consistency proof, that PA is indeed consistent. But then the formal application of U_5 to $\neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)$ is intuitively justified: the truth of $\neg (Bew_{PA}(\ulcorner 0 = 1 \urcorner) \wedge Bew_{PA}(\ulcorner 0 = 1 \urcorner))$ holds indeed in virtue of the truth of $\neg Bew_{PA}(\ulcorner 0 = 1 \urcorner)$. So, our predicational theory of ground proves an intuitively true claim about partial ground that the operational theory does not.

Thus, we know that the predicational theory of ground is stronger than the operational theory of ground. But how strong is it exactly? Here is the answer:

Theorem 2.5.10. *PG is a proof-theoretically conservative extension of PT.*

Proof. The proof is similar to the proof of Proposition 2.5.2. Let $\underset{\cdot}{c}$ represent the complexity function and $\underset{\cdot}{<}$ represent the strictly-less-than relation on the natural numbers again. We define the translation function $\tau : \mathcal{L}_{Tr}^{\lhd} \to \mathcal{L}_{Tr}$ recursively by saying that:

(i) $\tau(\varphi) = \begin{cases} \varphi, & \text{if } \varphi \in \mathcal{L}_{Tr} \text{ atomic,} \\ Tr(s) \wedge Tr(t) \wedge \underset{\cdot}{c}(s) \underset{\cdot}{<} \underset{\cdot}{c}(t), & \text{if } \varphi = s \lhd t; \end{cases}$

(ii) $\tau(\neg\varphi) = \neg\tau(\varphi)$;

(iii) $\tau(\varphi \circ \psi) = \tau(\varphi) \circ \tau(\psi)$, for $\circ = \wedge, \vee$; and

(iv) $\tau(Qx\varphi) = Qx(\tau(\varphi))$, for $Q = \forall, \exists$.[24]

It is again easy to see that (a) $\tau(\varphi) = \varphi$, if $\varphi \in \mathcal{L}_{Tr}$. Next we check that (b) the translation preserves theoremhood from PG to PT, in the sense that for all $\varphi \in \mathcal{L}^{\lhd}_{Tr}$, if $\vdash_{PG} \varphi$, then $\vdash_{PT} \tau(\varphi)$. We prove this result by an induction on the length of derivations. All the arithmetic axioms and rules of PG and PT are the same and all the instances of the induction scheme over $\mathcal{L}^{\lhd}_{Tr}$ become instances of the induction scheme over $\mathcal{L}_{Tr}$ under τ. Thus, it suffices to show that the images of the ground-theoretic axioms are derivable in PT. Here we just show the claim for two cases:

- In the case of the axiom G_1, we get $\tau(\forall x \neg(x \lhd x)) = \forall x \neg(Tr(x) \wedge Tr(x) \wedge \underset{\cdot}{c}(x) \underset{\cdot}{<} \underset{\cdot}{c}(x))$. We know that $\vdash_{PA} \forall x(Sent(x) \rightarrow \neg(\underset{\cdot}{c}(x) \underset{\cdot}{<} \underset{\cdot}{c}(x)))$. Since $\vdash_{PT} \forall x(Tr(x) \rightarrow Sent(x))$, the claim follows by simple logic.
- In the case of the axiom U_1, we get that $\tau(\forall x(Tr(x) \rightarrow x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)) = \forall x(Tr(x) \rightarrow Tr(x) \wedge Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \wedge \underset{\cdot}{c}(x) \underset{\cdot}{<} \underset{\cdot}{c}(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x))$. Now, let x be a fresh variable and assume that $Tr(x)$ for a $\rightarrow$-Intro in PT followed by a generalization. We need to derive $Tr(x) \wedge Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \wedge \underset{\cdot}{c}(x) \underset{\cdot}{<} \underset{\cdot}{c}(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$. The first conjunct of the consequent is simply the assumption $Tr(x)$. By the axiom P_1, $\forall x(Tr(x) \leftrightarrow Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x))$, we can derive $Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$ from the first conjunct and thus we get the second conjunct. Finally, for the last conjunct of the consequent is derivable, note that $\vdash_{PT} \forall x(Tr(x) \rightarrow Sent(x))$ and thus $\vdash_{PT} Tr(x) \rightarrow Sent(x)$, as well as $\vdash_{PT} \forall x(Sent(x) \rightarrow (\underset{\cdot}{c}(x) \underset{\cdot}{<} \underset{\cdot}{c}(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x))$. Thus, we get the third conjunct $\underset{\cdot}{c}(x) \underset{\cdot}{<} \underset{\cdot}{c}(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$ by simple logic. The claim follows.

The other axioms can be derived in a similar way under τ. The theorem follows by putting (a) and (b) together. $\square$

$\square$

The theorem has two important immediate consequences:

Corollary 2.5.11. *The theory PG is proof-theoretically consistent.*

Corollary 2.5.12. *The theory PG has the same arithmetic theorems as the theory ACA of arithmetical comprehension.*[25]

[24]Note that the translation of $s \lhd t$ is in the same spirit as the translation of $\varphi \prec \psi$ in the proof of Proposition 2.5.2: while there we had $\tau(\varphi \prec \psi) = \varphi \wedge \psi \wedge \underset{\cdot}{c}(\ulcorner\varphi\urcorner) \underset{\cdot}{<} \underset{\cdot}{c}(\ulcorner\psi\urcorner)$, we now have $\tau(s \lhd t) = Tr(s) \wedge Tr(t) \wedge \underset{\cdot}{c}(s) \underset{\cdot}{<} \underset{\cdot}{c}(t)$. In the case of $\varphi \prec \psi$ we didn't need to translate φ and ψ, since they were already assumed to be in $\mathcal{L}$. Similarly, here we know that if $\vdash_{PG} s \lhd t$, then by G_3 and T_3 we get that $\vdash_{PG} T(s) \wedge T(t)$ and $\vdash_{PG} Sent(s) \wedge Sent(t)$. In words, if $s \lhd t$ is provable in PG, then it's provable in PG that s and t are names of true sentences of $\mathcal{L}$. It is effectively this limitation of PG to partial ground between truths of the language of arithmetic that enables us to prove the conservativity result.

[25]For a definition of ACA, see [58, p. 107–8].

Proof. This follows from the facts that PG is a conservative extension of PT and that PT has the same arithmetical theorems as ACA.[26] □ □

Thus, Propositions 2.5.6 and Theorem 2.5.10 allow us to determine the proof-theoretic strength of our predicational theory of ground. Moreover, philosophically speaking, together the theorems shows that our predicational theory of ground and the theory of positive truth say the same things about truth. Since PG is a conservative extension of PT, the two theories prove exactly the same theorems in the language $\mathcal{L}_{Tr}$. Moreover, looking at the axioms of both theories, we can see that they paint the same truth-theoretic picture, as it were. Looking at the axioms of both theories, we see that they only contain *positive* occurrences of the truth predicate, where Tr occurs exclusively in the scope of an even number (in fact, zero) negations. In other words, according to both theories, the truths are built up successively from other, less complex truths—and never from falsehoods.[27] What the predicational theory of ground adds to this picture is that it stratifies the truths into a hierarchy according to their complexity: the result is the hierarchy of grounds. In this specific sense, the predicational framework allows us to make the truth-theoretic commitments of the theory of partial ground explicit.

2.5.3 Models for Partial Ground

In the last section, we have proved the consistency of the predicational theory of ground from the fact that it is a conservative extension of the consistent theory of positive truth. By the completeness theorem for first-order logic, we can infer from this that there is a first-order model of the predicational theory. But the way we proved this result does not give us any idea of what such a model looks like. In this subsection, we will construct a model for our theory "from scratch."

Consider the set $\mathbf{S}$ that contains all and only the codes of formulas that are true in the standard model of arithmetic:

$$\mathbf{S} =_{def} \{\#\varphi \mid \varphi \in \mathcal{L}, \mathbf{N} \vDash \varphi\}.$$

[26] We can prove this result via the equivalence of PT and CT (see Fn 22 and the usual proof that CT has the same arithmetic theorems as ACA. For the details of this proof, see [58, p. 101–16 respectively].

[27] For this point it matters that we are talking about PT and not the equivalent theory CT. Since CT has the axiom $\forall x(Tr(\dot{\neg}x) \leftrightarrow \neg Tr(x))$, where Tr occurs *negatively* in the scope of a single negation, the theory paints a different truth-theoretic picture. According to CT, the truths are build up from less complex truths *and falsehoods*. The point here is that the truth-theoretic picture painted by a theory is highly sensitive to the concrete axiomatization of the theory: even though CT and PT are equivalent, they paint a different picture.

Then $(\mathbf{N}, \mathbf{S})$ is a model of the theory of positive truth—the *standard model* of PT.[28] To construct our model, we will use *grounding-trees* over the standard model of PT. Grounding-trees were first introduced by Correia [25].[29] Here we give a slightly different definition of grounding-trees, which is adapted to the present purpose:

Definition 2.5.13. *Let* $(\mathbf{N}, \mathbf{S})$ *be the standard model of* PT*. We define the grounding-trees over* $(\mathbf{N}, \mathbf{S})$ *recursively by saying that for all formulas* $\varphi, \psi \in \mathcal{L}$*:*

(i) if $\#\varphi \in \mathbf{S}$*, then* $\#\varphi$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\varphi$ *as its root;*

(ii) if $\begin{array}{c}\#\varphi\\ \widehat{\mathcal{T}}\end{array}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\varphi$ *as its root, then* $\begin{array}{c}\#\neg\neg\varphi\\ | \\ \#\varphi\\ \widehat{\mathcal{T}}\end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg\neg\varphi$ *as its root;*

(iii) if $\begin{array}{c}\#\varphi\\ \widehat{\mathcal{T}}\end{array}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\varphi$ *as its root, then* $\begin{array}{c}\#(\varphi \vee \psi)\\ | \\ \#\varphi\\ \widehat{\mathcal{T}}\end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#(\varphi \vee \psi)$ *as its root;*

(iv) if $\begin{array}{c}\#\psi\\ \widehat{\mathcal{T}}\end{array}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\psi$ *as its root, then* $\begin{array}{c}\#(\varphi \vee \psi)\\ | \\ \#\psi\\ \widehat{\mathcal{T}}\end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#(\varphi \vee \psi)$ *as its root;*

(v) if $\begin{array}{c}\#\varphi\\ \widehat{\mathcal{T}_1}\end{array}, \begin{array}{c}\#\psi\\ \widehat{\mathcal{T}_2}\end{array}$ *are grounding-trees* $\mathcal{T}_1, \mathcal{T}_2$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\varphi, \#\psi$ *as their*

[28] For a proof, see [58, p. 116–22].

[29] In forthcoming work, Litland [84] and deRosset [31] argue for the ground-theoretic relevance of trees.

roots respectively, then $\begin{array}{c} \#(\varphi \wedge \psi) \\ \#\varphi \qquad \#\psi \\ \widehat{\mathcal{T}_1} \qquad \widehat{\mathcal{T}_2} \end{array}$ is a grounding-tree over $(\mathbf{N}, \mathbf{S})$ with $\#(\varphi \wedge \psi)$ as its root;

(vi) *if* $\begin{array}{c} \#\varphi(t) \\ \widehat{\mathcal{T}} \end{array}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\varphi(t)$ *as its root, then* $\begin{array}{c} \#\exists x \varphi(x) \\ | \\ \#\varphi(t) \\ \widehat{\mathcal{T}} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\exists x \varphi(x)$ *as its root;*

(vii) *if* $\begin{array}{c} \#\varphi(t_1) \\ \widehat{\mathcal{T}_1} \end{array}$, $\begin{array}{c} \#\varphi(t_2) \\ \widehat{\mathcal{T}_2} \end{array}$, $\ldots$ *are grounding-trees* $\mathcal{T}_1, \mathcal{T}_2, \ldots$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\varphi(t_1), \#\varphi(t_2), \ldots$ *as their roots respectively, where* $t_1, t_2, \ldots$ *are all and only the terms of* $\mathcal{L}_{PA}$*, then* $\begin{array}{c} \#\forall x \varphi(x) \\ \#\varphi(t_1) \qquad \#\varphi(t_2) \qquad \ldots \\ \widehat{\mathcal{T}_1} \qquad \widehat{\mathcal{T}_2} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\forall x \varphi(x)$ *as its root;*

(viii) *if* $\begin{array}{c} \#\neg\varphi \\ \widehat{\mathcal{T}} \end{array}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg\varphi$ *as its root, then* $\begin{array}{c} \#\neg(\varphi \wedge \psi) \\ | \\ \#\neg\varphi \\ \widehat{\mathcal{T}} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg(\varphi \wedge \psi)$ *as its root;*

(ix) *if* $\begin{array}{c} \#\neg\psi \\ \widehat{\mathcal{T}} \end{array}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg\psi$ *as its root, then* $\begin{array}{c} \#\neg(\varphi \wedge \psi) \\ | \\ \#\neg\psi \\ \widehat{\mathcal{T}} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg(\varphi \wedge \psi)$ *as its root;*

(x) *if* $\begin{array}{c} \#\neg\varphi \\ \widehat{\mathcal{T}_1} \end{array}$, $\begin{array}{c} \#\neg\psi \\ \widehat{\mathcal{T}_2} \end{array}$ *are grounding-trees* $\mathcal{T}_1, \mathcal{T}_2$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg\varphi, \#\neg\psi$

as their roots respectively, then $\begin{array}{c} \#\neg(\varphi \vee \psi) \\ \diagup \quad \diagdown \\ \underset{\widehat{\mathcal{T}_1}}{\#\neg\varphi} \quad \underset{\widehat{\mathcal{T}_2}}{\#\neg\psi} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg(\varphi \vee \psi)$ *as its root;*

(xi) if $\underset{\widehat{\mathcal{T}}}{\#\neg\varphi(t)}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg\varphi(t)$ *as its root, then* $\begin{array}{c} \#\neg\forall x\varphi(x) \\ \mid \\ \underset{\widehat{\mathcal{T}}}{\#\neg\varphi(t)} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg\forall x\varphi(x)$ *as its root;*

(xii) if $\underset{\widehat{\mathcal{T}_1}}{\#\neg\varphi(t_1)}$, $\underset{\widehat{\mathcal{T}_2}}{\#\neg\varphi(t_2)}$, ... *are grounding-trees* $\mathcal{T}_1, \mathcal{T}_2, \ldots$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\neg\varphi(t_1), \#\neg\varphi(t_2), \ldots$ *as their roots respectively, where* $t_1, t_2, \ldots$ *are all and only the terms of* $\mathcal{L}_{PA}$*, then* $\begin{array}{c} \#\neg\exists x\varphi(x) \\ \diagup \quad \mid \quad \diagdown \\ \underset{\widehat{\mathcal{T}_1}}{\#\neg\varphi(t_1)} \quad \underset{\widehat{\mathcal{T}_2}}{\#\neg\varphi(t_2)} \quad \ldots \end{array}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\forall x\varphi(x)$ *as its root;*

(xiii) nothing else is a grounding-tree over $(\mathbf{N}, \mathbf{S})$.

Mathematically speaking, the grounding-trees over $(\mathbf{N}, \mathbf{S})$ are *rooted graphs*, where the vertices are codes of formulas, one vertex is distinguished as the root, and the edges are constructed according to the above definition. Note that by clauses (vii) and (xii), there are infinitely *wide* grounding-trees over $(\mathbf{N}, \mathbf{S})$. Nevertheless, all grounding-trees over $(\mathbf{N}, \mathbf{S})$ have a finite *height*, where this concept is defined recursively on the construction of the grounding-trees over $(\mathbf{N}, \mathbf{S})$:

Definition 2.5.14. *We define the height* $h(\mathcal{T})$ *of a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, \mathbf{S})$ *by saying that:*

(i) all grounding-trees over $(\mathbf{N}, \mathbf{S})$ *of the form* $\#\varphi \in \mathbf{S}$ *have height one;*

(ii) if $\mathcal{T}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *that is constructed from grounding-trees* $\mathcal{T}_1, \mathcal{T}_2, \ldots$ *over* $(\mathbf{N}, \mathbf{S})$ *according to clauses (ii–xii) of Definition 2.5.13, then the height of* $\mathcal{T}$ *is one plus the least upper bound of the heights of* $\mathcal{T}_1, \mathcal{T}_2, \ldots$*:*

$$h(\mathcal{T}) = lub\{h(\mathcal{T}_1), h(\mathcal{T}_2), \ldots\} + 1,$$

where lub is the operation of taking least upper bounds.

We call a grounding-tree over $(\mathbf{N}, \mathbf{S})$ *degenerate iff it is of height one.*

The argument that grounding-trees have finite height now is a simple induction on the construction of the grounding-trees, where we note that in clauses (vii) and (xii) the height of the grounding-trees $\mathcal{T}_1, \mathcal{T}_2, \ldots$ is bounded by the (finite) logical complexity of φ. Thus, we can use *induction on the height of the grounding-trees* as a proof-method for claims about grounding-trees: if all degenerate grounding-trees have a property and we can show for all grounding-trees $\mathcal{T}$ that if all grounding-trees of lower height have the property, then $\mathcal{T}$ has the property, then *all* grounding-trees have the property.

For example, we can use this method to establish that the grounding-trees over $(\mathbf{N}, \mathbf{S})$ are really trees on the truths in $(\mathbf{N}, \mathbf{S})$:

Lemma 2.5.15. *Let* $(\mathbf{N}, \mathbf{S})$ *be the standard model of PT and let* $\mathcal{T}$ *be a grounding-tree over* $(\mathbf{N}, \mathbf{S})$*. Then for all formulas* $\varphi \in \mathcal{L}$*, if* $\#\varphi$ *is a vertex in* $\mathcal{T}$*, then* $\#\varphi \in \mathbf{S}$.

Proof. The claim trivially holds for all degenerate grounding-trees over $(\mathbf{N}, \mathbf{S})$ by clause (i) of Definition 2.5.13. So assume the induction hypothesis. For the induction step, we go through all the ways in which $\mathcal{T}$ could have been constructed according to Definition 2.5.13 and check that the claim holds. Here we only show this for clause (ii). Assume that $\mathcal{T}$ is of the form

$$\begin{array}{c} \#\neg\neg\varphi \\ | \\ \#\varphi \\ \widehat{\mathcal{T}'} \end{array}$$

, where $\overset{\#\varphi}{\widehat{\mathcal{T}'}}$ is a grounding-tree $\mathcal{T}'$ over $(\mathbf{N}, \mathbf{S})$ with $\#\varphi$ as its root.

Then, since $\mathcal{T}'$ is a grounding-tree over $(\mathbf{N}, \mathbf{S})$ with strictly lower height than $\mathcal{T}$, we know by the induction hypothesis that $\#\varphi \in \mathbf{S}$ and thus $\mathbf{N} \vDash \varphi$. But then, by classical logic, also $\mathbf{N} \vDash \neg\neg\varphi$ and thus $\#\neg\neg\varphi \in \mathbf{S}$. The claim follows. The cases for the other clauses are analogous. □

□

Next, we wish to show that the grounding-trees over $(\mathbf{N}, \mathbf{S})$ don't contain any cycles or "loops." This follows from the following useful lemma. Let us define the notion of an occurrence of the code $\#\varphi$ of a formula $\varphi \in \mathcal{L}$ to be *below* an occurrence of the code $\#\psi$ of formula $\psi \in \mathcal{L}$ in a grounding-tree $\mathcal{T}$ over $(\mathbf{N}, \mathbf{S})$ recursively by saying that: no occurrence of any formula in a degenerate grounding-tree over $(\mathbf{N}, \mathbf{S})$ is below an occurrence of any formula in the tree, and if $\mathcal{T}$ is a grounding-tree over $(\mathbf{N}, \mathbf{S})$ that was constructed from grounding-trees $\mathcal{T}_1, \mathcal{T}_2, \ldots$ over $(\mathbf{N}, \mathbf{S})$ according to the rules (ii–xii) of Definition 2.5.13, then all occurrences of all formulas in $\mathcal{T}_1, \mathcal{T}_2, \ldots$ occur below the root of $\mathcal{T}$ in $\mathcal{T}$. Then we can show:

Lemma 2.5.16. *Let* $(\mathbf{N}, \mathbf{S})$ *be the standard model of PT. If* $\mathcal{T}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\varphi$ *as its root, for some formula* φ*, then, all formulas* ψ *whose code* $\#\psi$ *occurs below* $\#\varphi$ *in* $\mathcal{T}$ *have a lower complexity than* φ*.*

Proof. The claim trivially holds for degenerate grounding-trees, since no code occurs below any code in degenerate grounding-trees. So assume the induction hypothesis. For the induction step, we note that in all the ways in which a grounding-tree can be constructed from other grounding-trees, the formula whose code is at the root of the new grounding-tree is the only new formula and always has higher complexity than the roots of the grounding-trees that it is constructed from. The claim follows than by a simple application of the induction hypothesis. □

□

Using this lemma, we obtain:
Lemma 2.5.17. *Let* $(\mathbf{N}, \mathbf{S})$ *be the standard model of PT and* $\mathcal{T}$ *is a grounding-tree over* $(\mathbf{N}, \mathbf{S})$*. Then between any two nodes* $\#\varphi$ *and* $\#\psi$ *in* $\mathcal{T}$*, for formulas* $\varphi, \psi \in \mathcal{L}$*, there is exactly one path, i.e. there are no cycles.*

Proof. Degenerate grounding-trees clearly contain no cycles. So assume the induction hypothesis. For the induction step, consider some arbitrary grounding-tree $\mathcal{T}$ with root $\#\varphi$, for some formula φ. By the induction hypothesis, we know that all the trees that $\mathcal{T}$ is constructed from don't contain any loops. Now assume that in the last step of the construction of $\mathcal{T}$ a loop is introduced. This would mean that $\#\varphi$ has to occur somewhere in the grounding-trees that $\mathcal{T}$ was constructed from. This would mean that $\#\varphi$ occurs below $\#\varphi$ in $\mathcal{T}$. But then by Lemma 2.5.16, we would get that $|\varphi| < |\varphi|$, which is impossible. The claim follows. □

□

Remember that a *rooted tree* in the mathematical sense is a rooted graph that does not contain any cycles. Thus, the grounding-trees over $(\mathbf{N}, \mathbf{S})$ are simply rooted trees over $\mathbf{S}$. In a rooted tree, all edges have a natural direction, either *towards* or *away* from the root. Given Lemma 2.5.16, we can say that all the edges in a grounding-tree $\mathcal{T}$ over $(\mathbf{N}, \mathbf{S})$ naturally point toward its root. Finally, we wish to show that the set of grounding-trees over $(\mathbf{N}, \mathbf{S})$ is *transitive* in the following sense:
Lemma 2.5.18. *Let* $(\mathbf{N}, \mathbf{S})$ *be the standard model of PT. If there is a grounding-tree* $\mathcal{T}_1$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\psi$ *as its root and* $\#\varphi_1, \#\varphi_2, \ldots$ *as its leaves and there is a grounding-tree* $\mathcal{T}_2$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\psi, \#\psi_1, \#\psi_2, \ldots$ *as its leaves and* $\#\theta$ *as its root, then there is a grounding-tree* $\mathcal{T}_3$ *over* $(\mathbf{N}, \mathbf{S})$ *with* $\#\varphi_1, \#\varphi_2, \ldots, \#\psi_1, \#\psi_2, \ldots$ *as its leaves and* $\#\theta$ *as its root.*

Proof. The proof proceeds by induction on the height of $\mathcal{T}_2$. If $\mathcal{T}_2$ is a degenerate grounding-tree over $(\mathbf{N}, \mathbf{S})$, then $\mathcal{T}_1$ is already the grounding-tree we are looking for. So, assume the induction hypothesis and let $\mathcal{T}_1$ and $\mathcal{T}_2$ be grounding-trees as described in the statement of the proposition. Now, we go through the different ways that $\mathcal{T}_2$ could have been constructed from other grounding-trees. Assume, for example, that $\mathcal{T}_2$ is constructed using clause (ii) of Definition 2.5.13. Then $\mathcal{T}_2$ is of the form: $\begin{array}{c} \#\neg\neg\theta' \\ | \\ \#\theta' \\ \widehat{} \\ \mathcal{T}_2' \end{array}$ and its root is $\#\theta = \#\neg\neg\theta'$. Since $\begin{array}{c} \#\theta' \\ \widehat{} \\ \mathcal{T}_2' \end{array}$ is then a grounding-tree with a strictly lower height than $\mathcal{T}_2$, by the induction hypothesis we get a grounding-tree $\mathcal{T}_3'$ with root $\#\theta'$ and $\#\varphi_1, \#\varphi_2, \dots, \#\psi_1, \#\psi_2, \dots$ as its leaves. Then by a simple application of (ii) of Definition 2.5.13, we get the existence of the desired grounding-tree $\mathcal{T}_3$ with root $\#\theta = \#\neg\neg\theta'$ and leaves $\#\varphi_1, \#\varphi_2, \dots, \#\psi_1, \#\psi_2, \dots$. The cases for the other clauses are analogous. □

□

The proof essentially shows that matching grounding-trees can be "glued together," as it were.

Putting all of the above together, we can say that the non-degenerate grounding-trees over $(\mathbf{N}, \mathbf{S})$ intuitively correspond to grounding-facts in $(\mathbf{N}, \mathbf{S})$: the codes at the leaves correspond to the truths that ground and the code at the root corresponds to the truth being grounded. More formally, Lemmas 2.5.17 and 2.5.18 together allow us to define a strict partial order on the elements of $\mathbf{S}$ (i.e. the truths according to $(\mathbf{N}, \mathbf{S})$): we say that a number $n \in \mathbf{S}$ is "strictly below" a number $m \in \mathbf{S}$ iff there is a non-degenerate grounding-tree connecting the two:

Definition 2.5.19. *Let* $(\mathbf{N}, \mathbf{S})$ *be the standard model of PT. We define the relation* $\mathbf{R} \subseteq \mathbb{N}^2$ *by saying that for all* $n, m \in \mathbb{N}$, $\mathbf{R}(m, n)$ *iff there is a non-degenerate grounding-tree over* $(\mathbf{N}, \mathbf{S})$ *with* n *as a leaf and* m *as its root.*

To be perfectly explicit, let's look at how $\mathbf{R}$ interprets the predicate $\lhd$ in the model $(\mathbf{N}, \mathbf{S}, \mathbf{R})$. Consider an atomic formula of the form $s \lhd t$ and a valuation σ in $(\mathbf{N}, \mathbf{S}, \mathbf{R})$. Using standard first-order semantics, we get: $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash_\sigma s \lhd t$ iff $\mathbf{R}(\llbracket s \rrbracket_\sigma, \llbracket t \rrbracket_\sigma)$, where $\llbracket t \rrbracket_\sigma$ is the value of the term t under the assignment σ.[30] In words, a formula $s \lhd t$ is true in $(\mathbf{N}, \mathbf{S}, \mathbf{R})$

[30]The value of a term under an assignment is defined in the usual recursive way. In the following, we only need that the value $\llbracket x \rrbracket_\sigma$ of a variable x under an assignment σ simply

under an assignment σ just in case there is a non-degenerate grounding-tree over $(\mathbf{N}, \mathbf{S})$ with the value of t under σ as its root and the value of s under σ as one of its leaves.

The relation $\mathbf{R}$ of Definition 6, then, indeed interprets the ground predicate correctly (in the way just described):

Theorem 2.5.20. $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash PG$.

Proof. We need to show that all the axioms of PG are satisfied in $(\mathbf{N}, \mathbf{S}, \mathbf{R})$. Since the basic truth axioms T_1, T_2, and T_3 are also axioms of PT, we know already that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \mathrm{T}_{1,2,3}$ as $(\mathbf{N}, \mathbf{S})$ is the standard model of PT. For the basic ground axioms G_1, G_2, and G_3 we get the following arguments:

- $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x \neg(x \lhd x)$

  This follows from Lemma 2.5.17: Assume that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \exists x(x \lhd x)$. Thus for some variable assignment σ over $(\mathbf{N}, \mathbf{S}, \mathbf{R})$ and some x-variant σ' of σ,[31] we have $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash_{\sigma'} x \lhd x$ meaning $\mathbf{R}(\sigma'(x), \sigma'(x))$. By Definition 2.5.19 this means that there is a non-degenerate grounding tree over $(\mathbf{N}, \mathbf{S})$ with $\sigma'(x)$ as its root and $\sigma'(x)$ as a leaf. But this tree would contain a loop, i.e. a path from $\sigma'(x)$ to $\sigma'(x)$, in contradiction to Lemma 2.5.17. Thus $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \nvDash \exists x(x \lhd x)$ and thus by classical logic $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x \neg(x \lhd x)$.

- $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x \forall y \forall z(x \lhd y \wedge y \lhd z \rightarrow x \lhd z)$

  This follows from Lemma 2.5.18: Let σ be a variable assignment over $(\mathbf{N}, \mathbf{S}, \mathbf{R})$ and σ' an x, y, z-variant σ such that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash_{\sigma'} x \lhd y$ and $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash_{\sigma'} y \lhd z$. This means that $\mathbf{R}(\sigma'(x), \sigma'(y))$ and $\mathbf{R}(\sigma'(y), \sigma'(z))$. By Definition 2.5.19 this means that are non-degenerate grounding-trees $\mathcal{T}_1$ and $\mathcal{T}_2$ over $(\mathbf{N}, \mathbf{S})$ such that $\mathcal{T}_1$ has $\sigma'(x)$ as a leaf and $\sigma'(y)$ as its root and $\mathcal{T}_2$ has $\sigma'(y)$ as a leaf and $\sigma'(z)$ as its root. By Lemma 2.5.18, there is a grounding-tree $\mathcal{T}_3$ over $(\mathbf{N}, \mathbf{S})$ with $\sigma'(x)$ as a leaf and $\sigma'(z)$ as its root. Thus, $\mathbf{R}(\sigma'(x), \sigma'(z))$ and thus $\vDash_{\sigma'} x \lhd z$. Since σ was arbitrary, we get that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x \forall y \forall z(x \lhd y \wedge y \lhd z \rightarrow x \lhd z)$.

- $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x \forall y(x \lhd y \rightarrow Tr(x) \wedge Tr(y))$

  This follows from Lemma 2.5.15: Let σ be a variable assignment over $(\mathbf{N}, \mathbf{S}, \mathbf{R})$ and σ' an x, y-variant of σ such that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash_{\sigma'} x \lhd y$. This means that $\mathbf{R}(\sigma'(x), \sigma'(y))$ and thus by Definition 2.5.19 that there is a non-degenerate grounding-tree over $(\mathbf{N}, \mathbf{S})$ with $\sigma'(x)$ as a leaf and $\sigma'(y)$ as its root. By Lemma 2.5.15 if follows that $\sigma'(x), \sigma'(y) \in \mathbf{S}$.

is $\sigma(x)$.

[31] Remember that a variable assignment σ' is an $\vec{x}$ variant of another variable assignment σ iff $\sigma'(\vec{x}) = \sigma(\vec{x})$.

Since $\mathbf{S}$ interprets Tr, we get that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash_{\sigma'} Tr(x) \wedge Tr(y)$. Since σ was arbitrary, we get $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x \forall y(x \lhd y \to Tr(x) \wedge Tr(y))$.

For the upward and downward directed axioms, we only consider the cases U_1 and D_1, as the other cases are analogous:

- $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x(Tr(x) \to x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$

  Let σ be a variable assignment over $(\mathbf{N}, \mathbf{S}, \mathbf{R})$ and σ' some x-variant of σ. Assume that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash_{\sigma'} Tr(x)$. This means that $\sigma'(x) \in \mathbf{S}$. Since $\mathbf{S} = \{\#\varphi \mid \varphi \in \mathcal{L}, \mathbf{N} \vDash \varphi\}$, we know that $\sigma'(x) = \#\varphi$, for some formula $\varphi \in \mathcal{L}$. Now, by clause (i) of Definition 2.5.13, $\#\varphi$ is a degenerate grounding-tree over $(\mathbf{N}, \mathbf{S})$. But then, by clause (ii) of Definition 2.5.13, $\begin{array}{c}\#\neg\neg\varphi \\ | \\ \#\varphi\end{array}$ is a non-degenerate grounding-tree over $(\mathbf{N}, \mathbf{S})$. Moreover, the root of this tree is $\#\neg\neg\varphi$ and its only leaf is $\#\varphi$. Now consider $\sigma'(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$. Since we know that $\sigma'(x) = \#\varphi$ and $\underset{\cdot}{\neg}$ expresses the negation function on the codes of formulas, we know that $\sigma'(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) = \#\neg\neg\varphi$. Thus, $\mathbf{R}(\sigma'(x), \sigma'(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x))$ meaning $\vDash_{\sigma'} x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x$. And since σ was arbitrary, we get $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x(Tr(x) \to x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$.

- $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \to x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$

  Let σ be a variable assignment over $(\mathbf{N}, \mathbf{S}, \mathbf{R})$ and σ' some x-variant of σ. Assume that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash_{\sigma'} Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$. This means that $\sigma'(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \in \mathbf{S}$. Again, since $\mathbf{S} = \{\#\varphi \mid \varphi \in \mathcal{L}, \mathbf{N} \vDash \varphi\}$ and since $\underset{\cdot}{\neg}$ represents the negation function on the codes of formulas, we know that $\sigma'(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) = \#\neg\neg\varphi$, for some formula $\varphi \in \mathcal{L}$ such that $\mathbf{N} \vDash \neg\neg\varphi$. From the latter it follows by classical logic that $\mathbf{N} \vDash \varphi$. Moreover, we know that $\sigma'(x) = \#\varphi$ and thus that $\#\sigma'(x) \in \mathbf{S}$. But then we get that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash_{\sigma'} Tr(x)$ and by the argument for the axiom U_1 that $\vDash_{\sigma'} x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x$. Hence, we get $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \forall x(Tr(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \to x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$, since σ was arbitrary. □

□

2.5.4 Grounding-Trees and Conceptualist Ground

We have shown how to extend the standard model of the theory of positive truth to a standard model of our predicational theory of ground. Note, however, that our construction will not necessarily work if we start from *non-standard* models of the theory of positive truth. At many points in our construction, we have made use of the fact that we're working in the standard model of the theory of positive truth. Most importantly, in our construction we have relied on the fact that the extension of the truth predicate coincides with the codes of formulas that are true in the standard model of PA. But in a non-standard model of the theory of positive truth this need not be the case. There we might have non-standard elements in the extension of the truth predicate, such as objects that according to the model are sentences with infinite complexity.[32] As a result, in such non-standard models, our construction will break down. Perhaps there is another construction that will allow us to extend non-standard models of the theory of positive truth to (non-standard) models of our predicational theory of ground. But here we leave this question open, as the point of the construction was to show what the standard model looks like.

By looking at the standard model of our predicational theory of truth, we can draw some lessons for the conceptualist semantics of ground. The model we have constructed is a model for a conceptualist notion of partial ground—it provides a semantics for our conceptualist ground predicate. Moreover, since we work in a standard first-order setting, we have a general notion of a model for predicational theories of ground. Developing a general conceptualist semantics for ground operators, in contrast, is still very much an open problem. But we can use the model that we have constructed to approach this problem. To begin with, we can interpret the operational theory of ground OG (Definition 2.5.1) over the model $(\mathbf{N}, \mathbf{S}, \mathbf{R})$. We simply say for all formulas $\varphi, \psi \in \mathcal{L}$:

$$(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \varphi \prec \psi \text{ iff}_{def} (\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \ulcorner \varphi \urcorner \lhd \ulcorner \psi \urcorner.$$

It then follows from Proposition 2.5.9 that this interpretation is sound with respect to OG, in the sense that for all $\varphi, \psi \in \mathcal{L}$, if $\vdash_{OG} \varphi \prec \psi$, then $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \varphi \prec \psi$.[33] For assume that $\vdash_{OG} \varphi \prec \psi$. By Proposition 2.5.9, it follows that $\vdash_{PG} \ulcorner \varphi \urcorner \lhd \ulcorner \psi \urcorner$. Then, by Theorem 2.5.20, we get that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \ulcorner \varphi \urcorner \lhd \ulcorner \psi \urcorner$ and thus by the above definition that $(\mathbf{N}, \mathbf{S}, \mathbf{R}) \vDash \varphi \prec \psi$. But this interpretation does not yet tell us what models for the language $\mathcal{L}_\prec$, in which the theory OG is formulated, should look like in general. But it gives us a hint about what kind of structure we can use to interpret the language. Remember that a *forest* in the mathematical sense is a *disjoint union* of trees. Now consider:

$$\mathbf{T} =_{def} \biguplus \{\mathcal{T} \mid \mathcal{T} \text{ is a non-degenerate grounding-tree over } (\mathbf{N}, \mathbf{S})\},$$

where $\uplus$ denotes the operation of taking disjoint unions. Then, mathematically speaking, $\mathbf{T}$ is a forest. As we have said before, the (non-degenerate) grounding-trees over $(\mathbf{N}, \mathbf{S})$ intuitively correspond to individual grounding facts. The forest $\mathbf{T}$, then, corresponds intuitively to the whole hierarchy of grounds. Moreover, we can equivalently rephrase our definition of $\mathbf{R}$ that

[32] For more on non-standard models of theories of syntax and truth, see [58, p. 83–89].

[33] Of course, the interpretation is not *complete*, in the sense that there are sentences $\varphi, \psi \in \mathcal{L}$ such that $\vDash \varphi \prec \psi$ and $\nvdash_{OG} \varphi \prec \psi$. This follows from the argument given above in §2.5.2 for the failure of the converse direction of Proposition 2.5.9.

interprets our ground predicate $\lhd$ (Definition 2.5.19) by saying that for all $m, n \in \mathbb{N}$:

$\mathbf{R}(m, n)$ iff there is a $\mathcal{T} \in \mathbf{T}$ with m as its leaf and n as its root.

With this definition in place, we can interpret *OG* over the structure $(\mathbf{N}, \mathbf{T})$. We simply equivalently restate the definition of $(\mathbf{N}, \mathbf{T}) \vDash \varphi \prec \psi$ by saying that for all $\varphi, \psi \in \mathcal{L}$:

$(\mathbf{N}, \mathbf{T}) \vDash \varphi \prec \psi$ iff there is a $\mathcal{T} \in \mathbf{T}$ with $\#\varphi$ as its leaf and $\#\psi$ as its root.

Now, take a look at the structure $(\mathbf{N}, \mathbf{T})$. On the above interpretation, the component $\mathbf{N}$ interprets the arithmetic vocabulary and in particular the names for sentences. The forest $\mathbf{T}$, on the other hand, interprets the ground operator $\prec$. By abstracting from this, we arrive at a notion of an arbitrary model for $\mathcal{L}_{\prec}$: The idea is that a model for $\mathcal{L}_{\prec}$ is a pair $(T, F(T))$, where T is a suitable set of fine-grained truths and $F(T)$ is a suitable forest over the elements of T. Of course, developing this idea in detail still requires a lot of work. But we wish to suggest that an intuitively plausible, *graph-theoretic semantics* for the ground operator can be obtained in this way. Incidentally, in forthcoming work, Litland [84] and deRosset [31] have already shown promising results in this direction—the results of this section give further support to this approach.

2.6 Partial Ground and Hierarchies of Typed Truth

In this section, we extend the present framework with principles for the grounds of truths that contain the truth predicate.

2.7 The Aristotelian Principle and Typed Truth

Aristotle has the following to say about the interaction of truth and explanation:

> *It is not because we think truly that you are pale, that you are pale; but because you are pale we who say this have the truth.*
> (Metaphysics 1051b6–9)

In the context of partial ground, this motivates the following principle: If φ is true, then it is natural to say that φ is true *in virtue of* what it says being the case, and if φ is false, then φ is false *in virtue of* what it says *not* being the case. We will thus call the corresponding informal ground-theoretic principles the *Aristotelian principles* about truth and falsehood respectively.

If we wish to formalize the Aristotelian principles in the context of our predicational theory of ground, we have to make a couple of adjustments. First, we have to assume that we have a Gödel-numbering for the language $\mathcal{L}_{Tr}$. In particular, we now assume that we have a name $\ulcorner \varphi \urcorner$ for every sentence $\varphi \in \mathcal{L}_{Tr}$. Second, we have to adjust our basic truth axioms. The axioms T_3 and G_3 of *PG* together ensure that both the truth predicate and the ground predicate only apply to sentences of $\mathcal{L}$. To formalize the Aristotelian principles, however, we need to relax this requirement: we need for the truth predicate and the ground predicate to apply to sentences of $\mathcal{L}_{Tr}$ that are not already in $\mathcal{L}$. To allow for this, we adjust the axiom T_3 to the axiom:

$$\forall x(Tr(x) \to Sent_{Tr}(x)),$$

which we'll label T_3^*. Together with G_3, this then entails:

$$\forall x \forall y(x \lhd y \to Sent_{Tr}(x) \land Sent_{Tr}(y)),$$

as well. We'll refer to the theory that results from replacing T_3 in *PG* with T_3^* as *PGT*. Thus, in *PGT*, we are not only talking about the truths of arithmetic, but also about the truth of the truths of arithmetic.

With these adjustments in place, we can schematically express the Aristotelian principle about truth by saying that for all sentences φ that:

$$\varphi \to \ulcorner \varphi \urcorner \lhd \ulcorner Tr(\ulcorner \varphi \urcorner) \urcorner.$$

For the Aristotelian principle about falsehood, we get:

$$\neg\varphi \to \ulcorner \neg\varphi \urcorner \lhd \ulcorner \neg Tr(\ulcorner \varphi \urcorner) \urcorner,$$

for all sentences φ. Using the same strategy as in the previous chapter, we can translate these schemata into the quantified axioms:

- (AP_T) $\quad \forall x(Tr(x) \to x \lhd \underset{\cdot}{Tr}(\dot{x}))$, and
- (AP_F) $\quad \forall x(Tr(\underset{\cdot}{\neg} x) \to \underset{\cdot}{\neg} x \lhd \underset{\cdot}{\neg} \underset{\cdot}{Tr}(\dot{x}))$,

where $\underset{\cdot}{Tr}$ represents the function that maps the code $\#t$ of a term t to the code $\#Tr(t)$ of the application $Tr(t)$ of the truth predicate to the term t. Thus, we have arrived at a quantified axiomatization of the Aristotelian principles.[34]

[34] To illustrate how the quantified axioms work, consider a formula φ such that $Tr(\ulcorner \varphi \urcorner)$. Then the principle AP_T says that $\ulcorner \varphi \urcorner \lhd \underset{\cdot}{Tr} \ulcorner \dot{\varphi} \urcorner$. But the latter is $\ulcorner \varphi \urcorner \lhd \ulcorner Tr(\ulcorner \varphi \urcorner) \urcorner$, since $\ulcorner \dot{\varphi} \urcorner = \ulcorner \ulcorner \varphi \urcorner \urcorner$ and $\underset{\cdot}{Tr}(\ulcorner \ulcorner \varphi \urcorner \urcorner) = \ulcorner Tr(\ulcorner \varphi \urcorner) \urcorner$. So AP_T allows us to prove $Tr(\ulcorner \varphi \urcorner) \to \ulcorner \varphi \urcorner \lhd \ulcorner Tr(\ulcorner \varphi \urcorner) \urcorner$, for every sentence φ. The quantified principle AP_F works analogously.

Unfortunately, as Fine [43] shows, the Aristotelian principles are *ground-theoretically inconsistent*:[35]

Lemma 2.7.1 (Puzzle of Ground)**.** *AP_T and AP_F are inconsistent over PGT.*

Proof. To show that $AP_T \vdash_{PGT} \bot$ consider the following derivation:

1. $\overline{0} = \overline{0}$ (Arithmetic)
2. $Tr(\ulcorner \overline{0} = \overline{0} \urcorner)$ (T-scheme over $\mathcal{L}$)
3. $\ulcorner \overline{0} = \overline{0} \urcorner \lhd \ulcorner Tr(\ulcorner \overline{0} = \overline{0} \urcorner) \urcorner$ (2., AP_T)
4. $Tr(\ulcorner Tr(\ulcorner \overline{0} = \overline{0} \urcorner) \urcorner)$ (3., G_3)
5. $\ulcorner Tr(\ulcorner \overline{0} = \overline{0} \urcorner) \urcorner \lhd \exists x Tr(x)$ (4., U_6)
6. $Tr(\ulcorner \exists x Tr(x) \urcorner)$ (5., G_3)
7. $\ulcorner \exists x Tr(x) \urcorner \lhd \ulcorner Tr(\ulcorner \exists x Tr(x) \urcorner) \urcorner$ (6., AP_T)
8. $Tr(\ulcorner Tr(\ulcorner \exists x Tr(x) \urcorner) \urcorner)$ (7., G_3)
9. $\ulcorner Tr(\ulcorner \exists x Tr(x) \urcorner) \urcorner \lhd \ulcorner \exists x Tr(x) \urcorner$ (8., U_6)
10. $\ulcorner \exists x Tr(x) \urcorner \lhd \ulcorner \exists x Tr(x) \urcorner$ (7.,9., G_2)
11. $\neg(\ulcorner \exists x Tr(x) \urcorner \lhd \ulcorner \exists x Tr(x) \urcorner)$ (G_1)
12. $\bot$ (10.,11., $\bot$-Intro)

To show that $AP_F \vdash_{PGT} \bot$, we can perform an analogous derivation, which is left to the interested reader. □ □

Note that the proof can easily be adapted to show that even the weaker, schematic formulations of the Aristotelian principles are ground-theoretically inconsistent.

We are left with a ground-theoretic puzzle about truth.[36] All the principles involved in the proof of Lemma 2.7.1 are intuitively plausible: the basic ground axioms G_2 and G_3 directly arise from the definition of partial ground, the upward directed axiom U_6 about the existential quantifier is plausible in light of the usual semantics for first-order logic, and the Aristotelian principle for truth is plausible from considerations about truth. (The principles

[35] Fine works in an operational framework, but his argument can easily be adapted to the present framework. Fine's argument is discussed and refined by Krämer [73] and Correia [25].

[36] Fine, in his original chapter [43], discusses a range of ground-theoretic puzzles that arise in a similar way from principles similar to the Aristotelian principle. Here we'll focus on the puzzle about truth and partial ground, because the problem arises most naturally in the present context.

required to show that AP_F is ground-theoretically inconsistent are equally plausible.) So what are we to do? In this section, we will propose a solution to the puzzle of ground that preserves the intuition behind all of these principles. We will achieve this by *typing* our truth predicate—a move familiar from *typed theories of truth*.

In typed theories of truth, no applications of the truth predicate to sentences containing the same truth predicate are provable. Thus, typed theories of truth respect Tarski's distinction between *object-language and meta-language* [132]. Tarski motivates this distinction from the semantic paradoxes, such as the infamous *liar paradox*. This paradox results when we apply the T-scheme to the *liar sentence* which intuitively "says of itself" that it is not true. More specifically, Tarski observed that, in a sufficiently strong background theory, such as PA, if we allow applications of the truth predicate to sentences with the same truth predicate in them, we get a sentence λ that is provably equivalent to its own falsehood:

$$\vdash_{PA} \lambda \leftrightarrow \neg Tr(\ulcorner \lambda \urcorner).$$

This follows from the so-called *diagonal lemma,* which is provable in PA in the context of an appropriate Gödel-numbering for $\mathcal{L}_{Tr}$:[37]

Lemma 2.7.2. *For every sentence $\varphi(x) \in \mathcal{L}_{Tr}$ with exactly one free variable x, there is a sentence $\delta \in \mathcal{L}_{Tr}$ such that*

$$\vdash_{PA} \delta \leftrightarrow \varphi(\ulcorner \delta \urcorner).$$

The existence of the liar sentence λ, then, follows by a simple application of the diagonal lemma to the formula $\neg Tr(x) \in \mathcal{L}_{Tr}$. It is well-known that the existence of a liar sentence is inconsistent with the T-scheme over the language $\mathcal{L}_{Tr}$.

A common intuitive response to the liar paradox is that it somehow arises from the self-reference involved. On this informal view, the problem is that the liar sentence "says something of itself," namely that it is not true.[38] Thus, so the intuitive response, we should put restrictions on our language that prevent self-reference. Tarski makes this response precise by introducing the distinction between object-language and meta-language. To illustrate the distinction, consider the truths of arithmetic. According to Tarski, if we wish to talk about numbers and their properties, we can do so in the language $\mathcal{L}$ of PA—our object language for arithmetic. But if we wish talk

[37]For a proof, see for example [16, p. 220–224].

[38]However, as Yablo [138] shows there are paradoxes without self-reference. Moreover, there are self-referential sentences that are not paradoxical. But still, the intuitive view is that in the case of the liar and similar paradoxes, self-reference plays an essential role.

about the truth of sentences in $\mathcal{L}$, we have to do so in the language $\mathcal{L}_{Tr}$—our meta-language for the truths of arithmetic.[39] Moreover, if we wish to talk about the truth of the truths of arithmetic, i.e. the truths of sentences in $\mathcal{L}_{Tr}$ containing the truth predicate, we need to do so in yet another *meta*-meta-language, which has a distinct truth-predicate for the sentences of $\mathcal{L}_{Tr}$. And so on. In contrast, Tarski calls a language that can talk about the truths of its own sentences, i.e. a language that has both names for all of its sentences and a truth predicate that applies to these names, *semantically closed*. Thus, a semantically closed language is its own meta-language, as it were, and thus we get self-referential paradoxes. Tarski shows that if we obey the distinction between object-language and meta-language, we can formulate a consistent theory of truth: In an appropriate meta-language, which is not semantically closed, we can consistently affirm the T-scheme for the sentences of the object-language and we never can prove problematic sentences, such as the liar. The liar paradox, on the other hand, shows that if we work in a semantically closed language, disaster ensues: If we have a Gödel-numbering for the terms of $\mathcal{L}_{Tr}$ within $\mathcal{L}_{Tr}$ and at the same time affirm the T-scheme over the sentences of $\mathcal{L}_{Tr}$, i.e. if we use $\mathcal{L}_{Tr}$ as its own meta-language, we get semantic paradoxes, like the liar paradox. Thus, so Tarski argues, when we wish to talk about truth, we should not never do so in a semantically closed language, but always in an appropriate meta-language. Intuitively, the picture is that semantic truths, such as truths about the truths of arithmetic, are on a "higher level" than non-semantic truths, such as the ordinary truths of arithmetic. Moreover, this can be iterated: the truths about truths about the truths of arithmetic are on yet a "higher level" than the truths about the truths of arithmetic and so on. What emerges is Tarski's *hierarchy of truths*. If we work in a semantically closed language, so Tarski, we mix the levels of the hierarchy of truths—and ultimately this is the source of the semantic paradoxes.[40]

Our original, unmodified predicational theory of ground PG respects Tarski's distinction between object- and meta-language: We have formulated PG in the language $\mathcal{L}^{\lhd}_{Tr}$ in the context of a coding for the language $\mathcal{L}$ of PA. In particular, we have assumed that we have a name $\ulcorner\varphi\urcorner$ for every sentence $\varphi \in \mathcal{L}$, but *not* that we have names $\ulcorner Tr(t)\urcorner$ for sentences of the form $Tr(t) \in \mathcal{L}_{Tr}$ and so on. Moreover, as we have said before, by the axioms T_3 and G_3, we have ensured that both the truth predicate and the ground predicate only apply to the sentences of $\mathcal{L}$. Thus, we have used the language $\mathcal{L}^{\lhd}_{Tr}$ as an appropriate meta-language for our object-language $\mathcal{L}$—in

[39]This language is then, of course, an *extension* of the language of PA: it extends the purely arithmetic vocabulary with names for the sentences of $\mathcal{L}$ and a truth predicate for those sentences.

[40]This also applies also to paradoxes without self-reference, such as Yablo's paradox: Yablo similarly formulates his paradox in a semantically closed language.

compliance with Tarski's distinction. When we move to the modified theory PGT, however, we no longer conform with Tarski's distinction. Since PGT is formulated in $\mathcal{L}_{Tr}^{\lhd}$ in the context of a Gödel numbering for $\mathcal{L}_{Tr}$, PGT is formulated in a semantically closed language. Now, the truth predicate may apply to sentences with the same truth predicate in them: Since in PGT we work in the context of a coding for $\mathcal{L}_{Tr}$, we have names for all the sentences of $\mathcal{L}_{Tr}^{\lhd}$ within $\mathcal{L}_{Tr}^{\lhd}$ itself. Moreover, by the axiom T_3^* we have allowed for these terms to occur truly in the context of the truth predicate and the ground predicate. In other words, when we formulated PGT, we have used $\mathcal{L}_{Tr}^{\lhd}$ as its own meta-language—we talked about the truths of $\mathcal{L}_{Tr}^{\lhd}$ within $\mathcal{L}_{Tr}^{\lhd}$ itself.

Based on this observation, we argue that the semantic closure of $\mathcal{L}_{Tr}^{\lhd}$ is (at least part of) the reason for the puzzle of ground. Note that the semantic closure of $\mathcal{L}_{Tr}^{\lhd}$ is required for the proof of Lemma 2.7.1. In the third step of the derivation, we applied the ground predicate to the truth predicate in $\ulcorner \overline{0} = \overline{0} \urcorner \lhd \ulcorner Tr(\ulcorner \overline{0} = \overline{0} \urcorner) \urcorner$. Moreover, in the fourth step, we inferred $Tr(\ulcorner Tr(\ulcorner \overline{0} = \overline{0} \urcorner) \urcorner)$ from this and thus applied the truth predicate to a sentence containing the same truth predicate. The main difference between the liar paradox and the puzzle of ground is that, in the case of the liar, we get a *truth-theoretic* inconsistency, i.e. an inconsistency with plausible principles for truth, while in the case of the puzzle, we get a *ground-theoretic* inconsistency, i.e. an inconsistency with plausible principles for partial ground. Still, the problematic sentences in both cases are quite similar. In both cases some intuitive form of self-reference is involved: while the liar sentence λ intuitively *says* something of itself, the principles of partial ground entail that truth of $\exists x Tr(x)$ partially *grounds* itself. Thus, we can say that the self-reference in the case of the liar is *semantic*, while the self-reference in the case of the puzzle of ground is *ground-theoretic*.[41]

In analogy to typed theories of truth, we propose a *typed solution* to the puzzle of ground. To formulate this solution, we will move to a slightly different framework, where instead of a single truth predicate Tr, we have a family $Tr_1, Tr_2, \ldots$ of *typed* truth predicates. These truth predicates intuitively express truth on the first, second, ... level of Tarski's hierarchy. In the remainder of this section, we will develop a *consistent* theory of partial ground and typed truth using these typed truth predicates. This theory will contain typed versions of the axioms of PG plus typed versions of the Aristotelian principles. Much like in the case of typed theories of truth, this will mean that no sentence is provable in which the truth predicate is applied to a sentence containing the same truth predicate. We will show that this

[41]There is also a kind of semantic self-reference involved in the case of puzzle. The existential quantifier in $\exists x Tr(x)$ semantically ranges over all sentences of $\mathcal{L}_{Tr}$, including $\exists x Tr(x)$ itself. Thus the truth of $\exists x Tr(x)$ is partially witnessed by the truth of $\exists x Tr(x)$. But here we do not wish to push this point any further.

restriction is sufficient to obtain a *consistent* theory of partial ground and typed truth.

2.8 Axiomatic Theories of Partial Ground and Typed Truth

Typed theories of truth aim to axiomatize Tarski's hierarchy of truths.[42] For this purpose, in typed theories of truth, we have different truth predicates for the different levels of the hierarchy. Correspondingly, we get a hierarchy of languages with a different language for every level of the hierarchy. To illustrate, we start with $\mathcal{L}_0 =_{def} \mathcal{L}$—the language of PA. The truth predicate for sentences of arithmetic is, then, Tr_1 and the language $\mathcal{L}_1$ extends $\mathcal{L}_0$ with Tr_1. The truth predicate for sentences of $\mathcal{L}_1$, in turn, is Tr_2 and the language $\mathcal{L}_2$ extends $\mathcal{L}_1$ with Tr_2. And so on. Thus, typed theories of truth are formulated using a hierarchical family of truth predicates $Tr_1, Tr_2, \ldots$ that intuitively correspond to truth on the different levels of Tarski's hierarchy.

2.8.1 Language and Background Theory

We will now formally define a hierarchy of languages, such that on every level we can talk about the truth of sentences on the lower levels. For reasons of generality, we will define this hierarchy in such a way that it includes even infinitary levels. Specifically, we assume that for every ordinal $0 < \alpha < \epsilon_0$, we have a different truth predicate T_α that intuitively expresses truth at the level α: we have $Tr_1, \ldots, Tr_\omega, \ldots, Tr_{\omega^\omega}, \ldots, Tr_{\omega^{\omega^\omega}}, \ldots$, where for $\alpha \neq \beta < \epsilon_0$, we have $Tr_\alpha \neq Tr_\beta$.[43] For all ordinals $1 \leq \alpha \leq \epsilon_0$, we define the language $\mathcal{L}_{<\alpha}$ as the language $\mathcal{L}$ of PA extended with all the truth predicates T_β for $0 < \beta < \alpha$:

$$\mathcal{L}_{<\alpha} =_{def} \mathcal{L} \cup \{Tr_\beta \mid 0 < \beta < \alpha\}.$$

Then we set:

$$\mathcal{L}_\alpha =_{def} \mathcal{L}_{<\alpha+1},$$

[42] For more on axiomatizations of Tarski's hierarchy, see [58, p. 125–29].

[43] We assume that the reader is familiar with the basic theory of ordinals. For the relevant definitions, see [65, p. 17–26], for example. The ordinal ϵ_0 is the limit of the sequence $1, \omega, \omega^\omega, \omega^{\omega^\omega}, \ldots$; in other words, ϵ_0 is the first ordinal that satisfies the equation $\omega^x = x$. This ordinal ϵ_0 is still *countable*, i.e. it has the same cardinality as the set of the natural numbers. But it provides a natural stopping point for our infinitary hierarchy, since (i) we can code the ordinals below ϵ_0 and (ii) PA proves the well-ordering of the ordinals below ϵ_0. We will not go into the details here, as the infinitary nature of our hierarchy is not particularly important to our *philosophical* point. Nevertheless, for reasons of generality, we will extend our hierarchy to this level, since ϵ_0 is the limit up to where we can apply the methods of this chapter.

for all ordinals $0 \leq \alpha < \epsilon_0$. Thus, the language $\mathcal{L}_0$ is $\mathcal{L}$, the language $\mathcal{L}_1$ is $\mathcal{L} \cup \{Tr_1\}$, and so on. Intuitively, for an ordinal $1 \leq \alpha \leq \epsilon_0$, the language $\mathcal{L}_{<\alpha}$ talks about the truths at the levels strictly below α and $\mathcal{L}_\alpha$ talks about the truths at all levels up to and including α. When we are operating on the ordinal level α, the language $\mathcal{L}_{<\alpha}$ is our intended object language, i.e. we wish to talk about grounding between the truths of sentences in $\mathcal{L}_{<\alpha}$. For most informal purposes, however, we already stop at the level of $\mathcal{L}_{<2} = \mathcal{L} \cup \{Tr_1\}$. The reason for this is that $\mathcal{L}_{<2}$ is the first language in which grounding between arithmetic truths and truths involving a truth predicate occurs. For all ordinals $1 \leq \alpha \leq \epsilon_0$, the language $\mathcal{L}^{\lhd}_{<\alpha}$ is $\mathcal{L}_{<\alpha}$ extended with our binary ground predicate $\lhd$:

$$\mathcal{L}^{\lhd}_{<\alpha} =_{def} \mathcal{L}_{<\alpha} \cup \{\lhd\}.$$

And we set:

$$\mathcal{L}^{\lhd}_{\alpha} =_{def} \mathcal{L}_{\alpha} \cup \{\lhd\},$$

for all ordinals $0 \leq \alpha < \epsilon_0$. When we are operating on the ordinal level α, we'll use $\mathcal{L}^{\lhd}_{\alpha}$ as our meta-language for the object-language language $\mathcal{L}_{<\alpha}$. Again, for expository purposes, we'll usually already stop at $\mathcal{L}^{\lhd}_{2} = \mathcal{L}_2 \cup \{\lhd\}$, which is the first language in which we can talk about grounding in $\mathcal{L}_{<2}$.

For an ordinal $0 < \alpha < \epsilon_0$, the theory $PAT_{<\alpha}$ is the result of extending PA with all the instances of the induction scheme over the language $\mathcal{L}_{<\alpha}$ and the theory $PAG_{<\alpha}$ is the result of extending $PAT_{<\alpha}$ with all the missing instances of the induction scheme over $\mathcal{L}^{\lhd}_{\alpha}$. For $0 \leq \alpha < \epsilon_0$, the theory PAT_α, then, is $PAT_{<\alpha+1}$ and similarly PAG_α is $PAG_{<\alpha+1}$. Thus, PAT_0 is PAT and PAG_0 is PAG. In PAT_α, we can develop a syntax theory for the languages $\mathcal{L}_{<\alpha}$ analogously to the way developed the syntax theory in the first part of this chapter. When we work on an ordinal level α, we assume that in $\mathcal{L}^{\lhd}_{\alpha}$, via some appropriate Gödel coding, we have names $\ulcorner\varphi\urcorner$ for all formulas $\varphi \in \mathcal{L}_{<\alpha}$.[44] Moreover, we assume that for every $0 < \beta < \alpha$, we have a function symbol $\dot{Tr_\beta}$ that represents the function which maps the code $\#t$ of a term t to the code $\#Tr_\beta(t)$ of the formula $Tr_\beta(t) \in \mathcal{L}_{<\alpha}$. And we abbreviate the formula that allows us to (strongly) represent the (set of codes of) sentences in $\mathcal{L}_{<\alpha}$ by $Sent_{<\alpha}$.

2.8.2 Axioms for Partial Ground and Typed Truth

With the syntax in place, we extend our theory PG to account for partial ground between truths on the same level of Tarski's hierarchy. We'll define this extension from the perspective of some ordinal level $0 < \alpha < \epsilon_0$. Thus,

[44] Here it is important that we're restricting ourselves to countable ordinals, because otherwise we would "run out of codes" at some point.

we wish to talk about ground between truths on all the ordinal levels $0 < \beta < \alpha$. To achieve this, we have to modify the basic ground axiom G_3 and the basic truth axioms T_1, T_2, and T_3. The axiom G_3 splits up in the following pair, for all ordinals $0 < \beta < \alpha$:

- (G^{β}_{3a}) $\quad \forall x \forall y(x \lhd y \rightarrow (Sent_{<\beta}(x) \rightarrow Tr_{\beta}(x)))$
- (G^{β}_{3b}) $\quad \forall x \forall y(x \lhd y \rightarrow (Sent_{<\beta}(y) \rightarrow Tr_{\beta}(y)))$

These axioms formalize the factivity of ground in a typed context. In particular, the axiom G^{β}_{3a} says that if the truth of some sentence grounds the truth of another, and if the sentence is below the level β in the hierarchy, then it is true at level β of the hierarchy. The axiom G^{β}_{3b}, on the other hand, says the same thing the other way around: if the truth of some sentence is grounded in the truth of another, and if the former sentence is below level β, then it is true at level β. To illustrate, if we let $\alpha = 2$, we get the following axiom pair:

- (G^{1}_{3a}) $\quad \forall x \forall y(x \lhd y \rightarrow (Sent_{<1}(x) \rightarrow Tr_{1}(x)))$
- (G^{1}_{3b}) $\quad \forall x \forall y(x \lhd y \rightarrow (Sent_{<1}(y) \rightarrow Tr_{1}(y)))$

Thus, for sentences $\varphi, \psi \in \mathcal{L}_{<1}$, the axioms G^{1}_{3a} and G^{1}_{3b} together say that the truth of φ can only ground the truth of ψ, if φ and ψ are both true at level one—i.e. if they are truths of arithmetic.

Next, in the typed versions of T_1 and T_2, we wish to make sure that true equations are true at every level below α. Thus, for all ordinals $0 < \beta < \alpha$, we postulate:

- (T^{β}_{1}) $\quad \forall s \forall t(Tr_{\beta}(s \dot{=} t) \leftrightarrow s^{\circ} = t^{\circ})$
- (T^{β}_{2}) $\quad \forall s \forall t(Tr_{\beta}(s \dot{\neq} t) \leftrightarrow s^{\circ} \neq t^{\circ})$

Thus, we get for example $Tr_1(\ulcorner \overline{0} = \overline{0} \urcorner), Tr_2(\ulcorner \overline{0} = \overline{0} \urcorner), \ldots$ and so on, for all ordinal levels below α.

In the case of T_3, we wish to make sure that a truth predicate Tr_{β}, for an ordinal $0 < \beta < \alpha$, only applies to sentences on levels below β—in compliance with Tarski's distinction. Thus, we postulate for all $0 < \beta < \alpha$:

- (T^{β}_{3}) $\quad \forall x(Tr_{\beta}(x) \rightarrow Sent_{<\beta}(x))$

Thus, for example, we get $\forall x(Tr_1(x) \rightarrow Sent_{<1}(x))$, which intuitively say that the predicate Tr_1 only applies to sentences of arithmetic.

Another adjustment is needed: Note that now $G^{\beta}_{3a/b}$ and T^{β}_{3} do *not* entail anymore that the ground predicate applies only to sentences. To ensure this, we postulate the following final basic ground axiom:

- (G^{α}_{4}) $\quad \forall x \forall y(x \lhd y \rightarrow Sent_{<\alpha}(x) \wedge Sent_{<\alpha}(y))$

Thus, on the level $\alpha = 2$, we get that $\forall x \forall y(x \lhd y \to Sent_{<2}(x) \wedge Sent_{<2}(y))$. In words: if $x \lhd y$, then both x and y are sentences of $\mathcal{L}_{<2}$, which is just the sentence of arithmetic $\mathcal{L}$ extended with the truth predicate Tr_1. Taken together, all of these modified axioms entail that our new theory respects Tarski's distinction between object- and meta-language.

Finally, we have to modify our upwards and downwards directed grounding axioms to apply on all levels below α. We achieve this by postulating that the axioms apply on all of these levels. Take the axioms U_1 and D_1 for example. They become the new typed set of axioms for all ordinals $0 < \beta < \alpha$:

- (U_1^β) $\quad \forall x(Tr_\beta(x) \to x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$
- (D_1^β) $\quad \forall x(Tr_\beta(\underset{\cdot}{\neg}\underset{\cdot}{\neg}x) \to x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$

Thus, on every level $0 < \beta < \alpha$, if a sentence is true on that level, then the sentence grounds its double-negation and if a double negation is true on the level, it is grounded by the sentence it is a double negation of.

Putting all of this together, we get:

Definition 2.8.1. *For all ordinals $0 \leq \alpha \leq \epsilon_0$, the predicational theory $PG_{<\alpha}$ of ground up to α, consists of the axioms of $PAG_{<\alpha}$ plus the following axioms for all $0 < \beta < \alpha$:*

Typed Ground Axioms:

G_1 $\forall x \neg(x \lhd x)$

G_2 $\forall x \forall y \forall z(x \lhd y \wedge y \lhd z \to x \lhd z)$

G_{3a}^β $\forall x \forall y(x \lhd y \to (Sent_{<\beta}(x) \to Tr_\beta(x)))$

G_{3b}^β $\forall x \forall y(x \lhd y \to (Sent_{<\beta}(y) \to Tr_\beta(y)))$

G_4^α $\forall x \forall y(x \lhd y \to Sent_{<\alpha}(x) \wedge Sent_{<\alpha}(y))$

Typed Truth Axioms:

T_1^β $\forall s \forall t(Tr_\beta(s \underset{\cdot}{=} t) \leftrightarrow s^\circ = t^\circ)$

T_2^β $\forall s \forall t(Tr_\beta(s \underset{\cdot}{\neq} t) \leftrightarrow s^\circ \neq t^\circ)$

T_3^β $\forall x(Tr_\beta(x) \to Sent_{<\beta}(x))$

Typed Upward Directed Axioms:

U_1^β $\forall x(Tr_\beta(x) \to x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg}x)$

U_2^β $\forall x \forall y(Tr_\beta(x) \to x \lhd x \underset{\cdot}{\vee} y \wedge Tr_\beta(y) \to y \lhd x \underset{\cdot}{\vee} y)$

U_3^β $\forall x \forall y(Tr_\beta(x) \wedge Tr_\beta(y) \to (x \lhd x \underset{\cdot}{\wedge} y) \wedge (y \lhd x \underset{\cdot}{\wedge} y))$

U_4^β $\forall x \forall y(Tr_\beta(\underset{\cdot}{\neg}x) \wedge Tr_\beta(\underset{\cdot}{\neg}y) \to (\underset{\cdot}{\neg}x \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)) \wedge (\underset{\cdot}{\neg}y \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)))$

U_5^β $\forall x \forall y(Tr_\beta(\underset{\cdot}{\neg}x) \to \underset{\cdot}{\neg}x \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\wedge} y) \wedge Tr_\beta(\underset{\cdot}{\neg}y) \to \underset{\cdot}{\neg}y \lhd \underset{\cdot}{\neg}(x \underset{\cdot}{\wedge} y))$

U_6^β $\forall x \forall t \forall v(Tr_\beta(x(t/v)) \to x(t/v) \lhd \underset{\cdot}{\exists}vx)$

U_7^β $\forall x \forall v(\forall t Tr_\beta(\underset{\cdot}{\neg}x(t/v)) \to \forall t(\underset{\cdot}{\neg}x(t/v) \lhd \underset{\cdot}{\neg}\underset{\cdot}{\exists}vx))$

U_8^β $\forall x \forall v(\forall t(Tr_\beta(x(t/v)) \to \forall t(x(t/v) \lhd \underset{\cdot}{\forall}vx))$

U_9^β $\forall x \forall t \forall v(Tr_\beta(\dot{\neg} x(t/v)) \to \dot{\neg} x(t/v) \lhd \dot{\neg}\dot{\forall} vx))$

Typed Downward Directed Axioms:

D_1 $\forall x(Tr_\beta(\dot{\neg}\dot{\neg} x) \to x \lhd \dot{\neg}\dot{\neg} x)$

D_2 $\forall x \forall y(Tr_\beta(x \dot{\vee} y) \to (Tr_\beta(x) \to x \lhd x \dot{\vee} y) \wedge (Tr_\beta(y) \to y \lhd x \dot{\vee} y))$

D_3 $\forall x \forall y(Tr_\beta(x \dot{\wedge} y) \to (x \lhd x \dot{\wedge} y) \wedge (y \lhd x \dot{\wedge} y))$

D_4 $\forall x \forall y(Tr_\beta(\dot{\neg}(x \dot{\wedge} y)) \to (Tr_\beta(\dot{\neg} x) \to \dot{\neg} x \lhd \dot{\neg}(x \dot{\vee} y)) \wedge (Tr_\beta(\dot{\neg} y) \to \dot{\neg} y \lhd \dot{\neg}(x \dot{\vee} y)))$

D_5 $\forall x \forall y(Tr_\beta(\dot{\neg}(x \dot{\vee} y)) \to (\dot{\neg} x \lhd \dot{\neg}(x \dot{\vee} y)) \wedge (\dot{\neg} y \lhd \dot{\neg}(x \dot{\vee} y)))$

D_6 $\forall x(Tr_\beta(\dot{\exists} vx(v)) \to \exists t(x(t/v) \lhd \dot{\exists} vx))$

D_7 $\forall x \forall v(Tr_\beta(\dot{\neg}\dot{\exists} vx) \to \forall t(\dot{\neg} x(t/v) \lhd \dot{\neg}\dot{\exists} vx))$

D_8 $\forall x \forall v(Tr_\beta(\dot{\forall} vx \to \forall t(x(t/v) \lhd \dot{\forall} vx))$

D_9 $\forall x \forall v(Tr_\beta(\dot{\neg}\dot{\forall} vx) \to \exists t(\dot{\neg} x(t/v) \lhd \dot{\neg}\dot{\forall} vx))$

For $0 \leq \alpha < \epsilon_0$, we define PG_α as $PG_{<\alpha+1}$.

To illustrate what PG_α looks like, for different α's, let's consider at a few examples. First, note that PG_0 is PAG. Next, note PG_1 is a functional analog of our original theory PG, where the truth-predicate has been "renamed" Tr_1. In particular, we get that PG_1 proves the theory PT of positive truth.

Since for all $1 < \alpha < \epsilon_0$, PG_α contains PG_1, we can say that PG_α essentially is (in the precise sense sketched above) an extension of PG. For α bigger than one, PG_α essentially consists of α-many copies of PG, one for every $\mathcal{L}_\beta$ and truth predicate Tr_β, where $1 < \beta < \alpha$. What is new in those theories is that now (names of) sentences involving the truth predicate may occur in the context of the ground predicate and other truth predicates—as long as we respect the typing restriction that for all $0 < \beta \leq \alpha$, if $Tr_\beta(\ulcorner\varphi\urcorner)$, then $Sent_{<\beta}(\ulcorner\varphi\urcorner)$. For example, in PG_2, we get the following instance of U_1^2:

$$Tr_2(\ulcorner Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner) \to \ulcorner Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner \lhd \ulcorner\neg\neg Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner.$$

Indeed, using G_{3b}^2 we can infer from this that:

$$\ulcorner Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner \lhd \ulcorner\neg\neg Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner \to Tr_2(\ulcorner\neg\neg Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner),$$

which together with the previous formula gives us:

$$Tr_2(\ulcorner Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner) \to Tr_2(\ulcorner\neg\neg Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner).$$

The other direction:

$$Tr_2(\ulcorner\neg\neg Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner) \to Tr_2(\ulcorner Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner)$$

can be shown analogously using D^2_1 and G^2_{3a}. Generalizing this idea, we get more substantial truth-theoretic theorems in PG_2, such as:

$$\forall x(Tr_2(\ulcorner Tr_1(\dot{x})\urcorner) \leftrightarrow Tr_2(\ulcorner\neg\neg Tr_1(\dot{x})\urcorner)),$$

for example. But so far, PG_2 does not allow us to prove any theorems of the form $Tr_2(\ulcorner Tr_1(\ulcorner\varphi\urcorner)\urcorner)$, where $\varphi \in \mathcal{L}$. In other words, we can't prove the truth of any sentence involving a truth predicate—even if they respect the typing restrictions. Thus, PG_2 is not really a theory of truth at level 2 yet—it can't even show that $Tr_2(\ulcorner Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner)$. Moreover, in PG_2, we can't prove any theorems of the form $\ulcorner\varphi\urcorner\lhd\ulcorner Tr_1(\ulcorner\varphi\urcorner)\urcorner$ or the like, where the ground predicate applies to sentence involving a truth predicate. To make things worse, all of this doesn't change on any level $\alpha > 2$. To get a more substantial theory of ground and partial truth, we need to say something about the grounds of truths involving the truth predicate: we need typed versions of the Aristotelian principles.

Typing the Aristotelian principles for use on an ordinal level α is pretty straight-forward. We get the following axioms for every $\gamma < \alpha$:

- (APU^γ_T) $\quad \forall x(Tr_\gamma(x) \to x \lhd \underset{\cdot}{Tr}_\gamma(\dot{x}))$
- (APU^γ_F) $\quad \forall x(Tr_\gamma(\underset{\cdot}{\neg}x) \to \underset{\cdot}{\neg}x \lhd \underset{\cdot}{\neg}\underset{\cdot}{Tr}_\gamma(\dot{x}))$

The axiom APU^1_T, for example, allows us to derive that $\ulcorner\overline{0} = \overline{0}\urcorner\lhd\ulcorner Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)\urcorner$ using the fact that by axiom T^1_1 we have $Tr_1(\ulcorner\overline{0} = \overline{0}\urcorner)$. The axioms $\mathrm{APU}^\beta_{T/F}$ are upwards directed axioms. For analogous reasons as in the case of the other ground axioms, we also need downward directed axioms for the Aristotelian principles. Again, straight-forwardly, we get for all $\gamma < \beta \leq \alpha$:

- ($\mathrm{APD}^{\beta;\gamma}_T$) $\quad \forall x(Tr_\beta(\underset{\cdot}{Tr}_\gamma(\dot{x})) \to x \lhd \underset{\cdot}{Tr}_\gamma(\dot{x}))$
- ($\mathrm{APD}^{\beta;\gamma}_F$) $\quad \forall x(Tr_\beta(\underset{\cdot}{\neg}\underset{\cdot}{Tr}_\gamma(\dot{x})) \to \underset{\cdot}{\neg}x \lhd \underset{\cdot}{\neg}\underset{\cdot}{Tr}_\gamma(\dot{x}))$

If we add the upward and downward directed versions of the Aristotelian principles to the previous theory, we arrive at our theory of ground and typed truth:

Definition 2.8.2. *For every ordinal $0 \leq \alpha < \epsilon_0$, the theory PGA_α of partial ground with the Aristotelian principles up (and including) to α consists of the axioms of PG_α plus the following axioms for all $\gamma < \beta \leq \alpha$:*

Upward Directed Aristotelian Principles:

- *(APU^γ_T)* $\quad \forall x(Tr_\gamma(x) \to x \lhd \underset{\cdot}{Tr}_\gamma(\dot{x}))$
- *(APU^γ_F)* $\quad \forall x(Tr_\gamma(\underset{\cdot}{\neg}x) \to \underset{\cdot}{\neg}x \lhd \underset{\cdot}{\neg}\underset{\cdot}{Tr}_\gamma(\dot{x}))$

Downward Directed Aristotelian Principles:

- *($APD^{\beta;\gamma}_T$)* $\quad \forall x(Tr_\beta(\underset{\cdot}{Tr}_\gamma(\dot{x})) \to x \lhd \underset{\cdot}{Tr}_\gamma(\dot{x}))$

- *($APD_F^{\beta;\gamma}$)* $\quad \forall x(Tr_\beta(\dot\neg Tr_\gamma(\dot x)) \to \dot\neg x \lhd \dot\neg Tr_\gamma(\dot x))$

The theory $PGA_{<\alpha}$ is defined as $\bigcup_{\beta<\alpha} PGA_\alpha$, for all $0 < \alpha \leq \epsilon_0$.

To see how PGA_α works for different $0 \leq \alpha < \epsilon_0$, let's consider again a few examples. First note that PGA_0 is PG_0 which is just PAG. Similarly, PGA_1 is PG_1. Things get interesting at the level PGA_2. Here we get:

$$Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner) \to \ulcorner\overline{0}=\overline{0}\urcorner \lhd \ulcorner Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner)\urcorner,$$

by instantiating the axiom APU_T^1 with the term $\ulcorner\overline{0}=\overline{0}\urcorner$. Moreover, by instantiating the axiom T_1^1 with the same term, we have:

$$Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner).$$

So putting the two together, we get:

$$\ulcorner\overline{0}=\overline{0}\urcorner \lhd \ulcorner Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner)\urcorner.$$

Now, using the instance:

$$\ulcorner\overline{0}=\overline{0}\urcorner \lhd \ulcorner Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner)\urcorner \to (Sent_{<2}(\ulcorner Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner)\urcorner) \to Tr_2(\ulcorner Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner)\urcorner))$$

of the axiom G_{3b}^2, and since:

$$Sent_{<2}(\ulcorner Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner)\urcorner)$$

is derivable in PA, we can infer:

$$Tr_2(\ulcorner Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner)\urcorner).$$

So, in PGA_2, we can indeed derive applications of the truth predicate to sentences with a truth predicate in them. Moreover, by putting APU_T^1:

$$\forall x(Tr_1(x) \to x \lhd T\dot{r}_1(\dot x))$$

and $APD_T^{2,1}$:

$$\forall x(Tr_2(T\dot{r}_1(\dot x)) \to x \lhd T\dot{r}_1(\dot x))$$

together, we can actually prove:

$$\forall x(Tr_2(T\dot{r}_1(\dot x)) \leftrightarrow Tr_1(x))$$

using the axioms $\mathrm{G}_{3a/b}^1$ and $\mathrm{T}_{1/2}^1$.[45] Thus, PGA_2 proves intuitive truths at level two, such as $Tr_2(\ulcorner Tr_1(\ulcorner\overline{0}=\overline{0}\urcorner)\urcorner)$, as well as quite substantial truth-theoretic principles, such as $\forall x(Tr_2(T\dot{r}_1(\dot x)) \leftrightarrow Tr_1(x))$. In other words, PG_2 proves something that looks like a substantial theory of truth at level two of Tarski's hierarchy. In the next section, we will show that for $0 < \alpha < \epsilon_0$, PGA_α proves the theory PRT_α of *positive ramified truth up to α*. Indeed, we can show that PGA_α is a *proof-theoretically conservative extension* of PRT_α.

[45] For the proof note that if we assume that $Tr_1(t)$ for a term t, it follows by axiom T_1^1 that $Sent_{<1}(t)$ and thus we can prove in PA that $Sent_{<2}(T\dot{r}_1(t))$. Similarly, if we assume $Tr_2(T\dot{r}_1\dot t)$, we can prove that $Sent_{<1}(t)$ by T_3^2 and PA.

2.8.3 Conservativity and Models

The theory $PT_{<\alpha}$ of positive ramified truth up to an ordinal level $1 \leq \alpha \leq \epsilon_0$ is formulated in the language $\mathcal{L}_{<\alpha}$ and it is the result of modifying the theory of typed truth with the typed versions of its axioms in a similar way as we developed $PG_{<\alpha}$:

Definition 2.8.3 ('Positive Ramified Truth')**.** *For all ordinals* $1 \leq \alpha \leq \epsilon_0$, *the theory* $PRT_{<\alpha}$ *of positive ramified truth up to* α *consists of the axioms of* $PAT_{<\alpha}$ *plus the following axioms for all* $\gamma < \beta < \alpha$*:*

Typed Truth Axioms:

T_1^β $\forall s \forall t(Tr_\beta(s \underset{\cdot}{=} t) \leftrightarrow s^\circ = t^\circ)$

T_2^β $\forall s \forall t(Tr_\beta(s \underset{\cdot}{\neq} t) \leftrightarrow s^\circ \neq t^\circ)$

T_3^β $\forall x(Tr_\beta(x) \to Sent_{<\beta}(x))$

Positive Ramified Truth Axioms:

RP_1^β $\forall x(Tr_\beta(x) \leftrightarrow Tr_\beta(\underset{\cdot}{\neg}\underset{\cdot}{\neg} x))$

RP_2^β $\forall x \forall y(Tr_\beta(x \underset{\cdot}{\wedge} y) \leftrightarrow Tr_\beta(x) \wedge Tr_\beta(y))$

RP_3^β $\forall x \forall y(Tr_\beta(\underset{\cdot}{\neg}(x \underset{\cdot}{\wedge} y)) \leftrightarrow Tr_\beta(\underset{\cdot}{\neg} x \vee T_\beta \underset{\cdot}{\neg} y))$

RP_4^β $\forall x \forall y(Tr_\beta(x) \underset{\cdot}{\vee} Tr_\beta(y) \leftrightarrow Tr_\beta(x) \vee Tr_\beta(y))$

RP_5^β $\forall x \forall y(Tr_\beta(\underset{\cdot}{\neg}(x \underset{\cdot}{\vee} y)) \leftrightarrow Tr_\beta(\underset{\cdot}{\neg} x) \wedge Tr_\beta(\underset{\cdot}{\neg} y))$

RP_6^β $\forall x \forall v(Tr_\beta(\underset{\cdot}{\forall} v x) \leftrightarrow \forall t Tr_\beta(x(t/v)))$

RP_7^β $\forall x \forall v(Tr_\beta(\underset{\cdot}{\neg} \underset{\cdot}{\forall} v x) \leftrightarrow \exists t Tr_\beta(\underset{\cdot}{\neg} x(t/v)))$

RP_8^β $\forall x \forall v(Tr_\beta(\underset{\cdot}{\exists} v x) \leftrightarrow \exists t Tr_\beta(x(t/v)))$

RP_9^β $\forall x \forall v(Tr_\beta(\underset{\cdot}{\neg} \underset{\cdot}{\exists} v x) \leftrightarrow \forall t Tr_\beta(\underset{\cdot}{\neg} x(t/v)))$

RP_{10}^β $\forall x(Tr_\beta(\underset{\cdot}{Tr_\gamma} \dot{x}) \leftrightarrow Tr_\gamma(x))$

RP_{11}^β $\forall x(Tr_\beta(\underset{\cdot}{\neg} \underset{\cdot}{Tr_\gamma} \dot{x}) \leftrightarrow Tr_\gamma(\underset{\cdot}{\neg} x))$

$RP_{12}^{\gamma,\beta}$ $\forall x(Sent_{<\gamma}(x) \to (Tr_\beta(\underset{\cdot}{Tr_\gamma} \dot{x}) \leftrightarrow Tr_\beta(x)))$

$RP_{13}^{\gamma,\beta}$ $\forall x(Sent_{<\gamma}(x) \to (Tr_\beta(\underset{\cdot}{\neg} \underset{\cdot}{Tr_\gamma} \dot{x}) \leftrightarrow Tr_\beta(\underset{\cdot}{\neg} x)))$

The theory PRT_α, *for* $0 \leq \alpha < \epsilon_0$, *is defined as* $RPT_{<\alpha+1}$.

Note that the theory PRT_1 is a functional analog of PT in the same way that PG_1 is a functional analog of PG. The theory $PRT_{<\alpha}$, for $1 < \alpha \leq \epsilon_0$, however, is a much stronger theory of truth than PT—it formalizes the

Tarskian hierarchy up to the level α.[46] For example, PGT_2 contains the axioms:

$$Tr_2(\ulcorner Tr_1(\ulcorner \overline{0} = \overline{0}\urcorner)\urcorner),$$

and

$$\forall x(Tr_2(T\underset{\cdot}{r}_1(\dot{x})) \leftrightarrow Tr_1(x)),$$

just like PGA_2. Indeed, we get:
Proposition 2.8.4. *For all ordinals* $1 \leq \alpha \leq \epsilon_0$, *the theory* $PGA_{<\alpha}$ *proves the theory* $PRT_{<\alpha}$*:* $PGA_{<\alpha} \vdash PRT_{<\alpha}$.

Proof. In large parts, the proof is analogous to the proof of Proposition 4.6 of the first part of this chapter: we carry out the same argument for all $\beta < \alpha$. The interesting cases are the new axioms RP^{β}_{10-13}, for $\beta < \alpha$, which can be shown from the typed Aristotelian principles $APU^{\gamma}_{T/F}$ and $APD^{\beta,\gamma}_{T/F}$, for $\gamma < \beta \leq \alpha$, the typed truth axiom T^{β}_{3}, and the typed ground axioms $G^{\beta}_{3a/b}$, for $\beta < \alpha$, and G^{α}_{4}. Here we only show how to derive RP^{β}_{10} and RP^{β}_{11}, as the other axioms are analogous:

- $\vdash_{PGA_{<\alpha}} \forall x(Tr_\beta(T\underset{\cdot}{r}_\gamma \dot{x}) \leftrightarrow Tr_\gamma(x))$ for $\gamma < \beta < \alpha$

  Let x be a fresh variable for $\forall$-Intro. We now prove both directions of the biconditional $Tr_\beta(T\underset{\cdot}{r}_\gamma \dot{x}) \leftrightarrow Tr_\gamma(x)$.

  "$\rightarrow$:" Assume $(*)$ $Tr_\beta(T\underset{\cdot}{r}_\gamma(x))$ for a $\rightarrow$-Intro. By T^{β}_{3}, we can derive $Sent_{<\beta}(T\underset{\cdot}{r}_\gamma(x))$. Using PA and T^{γ}_{3}, we can derive $(**)$ $Sent_{<\gamma}(x)$ from this. Moreover, using $(*)$ and $APD^{\beta,\gamma}_{T}$:

  $$\forall x(Tr_\beta(T\underset{\cdot}{r}_\gamma(\dot{x})) \rightarrow x \lhd T\underset{\cdot}{r}_\gamma(\dot{x})),$$

  we can derive $x \lhd T\underset{\cdot}{r}_\gamma(\dot{x})$. Using $(**)$ and G^{γ}_{4}:

  $$\forall x \forall y(x \lhd y \rightarrow (Sent_{<\gamma}(x) \rightarrow Tr_\gamma(x))),$$

  we can in turn derive: $Tr_\gamma(x)$. Thus, we get $Tr_\beta(T\underset{\cdot}{r}_\gamma(x)) \rightarrow Tr_\gamma(x)$ by $\rightarrow$-Intro.

[46] We can formulate a slightly stronger version of $PRT_{<\alpha}$ by replacing the schematic axioms RP^{β}_{12} and RP^{β}_{13} with the quantified axioms: $\forall x \forall \gamma \underset{\cdot}{<} \ulcorner \beta \urcorner (Sent_{<\gamma}(x) \rightarrow (Tr_\beta(T\underset{\cdot}{r}_\gamma \dot{x}) \leftrightarrow Tr_\beta(x)))$ and $\forall x \forall \gamma \underset{\cdot}{<} \ulcorner \beta \urcorner (Sent_{<\gamma}(x) \rightarrow (Tr_\beta(\underset{\cdot}{\neg} T\underset{\cdot}{r}_\gamma \dot{x}) \leftrightarrow Tr_\beta(\underset{\cdot}{\neg} x)))$, where $\forall \gamma$ quantifies over *codes of ordinals*, $\ulcorner \beta \urcorner$ is a term for a code of the ordinal β, and $\underset{\cdot}{<}$ represents the well-ordering on ordinals. To properly formulate these axioms, we require a coding of the ordinals up to ϵ_0, a representation of the natural well-ordering of these ordinals, and a justification for quantifying into ordinal indexes in $Sent_{<\gamma}$ and $T\underset{\cdot}{r}_\gamma$. We will discuss such a coding below, but for reasons of perspicuity, we will stick with the slightly weaker schematic version of $PRT_{<\alpha}$. Also, the (schematic) theory $PRT_{<\alpha}$ is equivalent to the (schematic) theory $RT_{<\alpha}$ of *ramified truth* up to α, which is the typed version of CT. As in the case of PT and CT, we will take the meta-theorems of $RT_{<\alpha}$ and apply them immediately to $PRT_{<\alpha}$.

"←:" Assume $(*')$ $Tr_\gamma(x)$ for another →-Intro. Using PA and G_4^γ, we get $Sent_{<\gamma}(x)$. From this and PA, we can derive for all $\gamma < \beta < \alpha$ that (**') $Sent_{<\beta}(\underset{\cdot}{Tr}_\gamma(\dot{x}))$. Moreover, using $(*')$ and APU_T^γ:

$$\forall x(Tr_\gamma(x) \rightarrow x \lhd \underset{\cdot}{Tr}_\gamma(\dot{x})),$$

we get $x \lhd \underset{\cdot}{Tr}_\gamma(\dot{x}))$. From this, using $(**')$ and G_{3b}^β, we get $Tr_\beta(\underset{\cdot}{Tr}_\gamma(x))$ and thus $Tr_\gamma(x) \rightarrow Tr_\beta(\underset{\cdot}{Tr}_\gamma(x))$ by →-Intro.

Putting both "→" and "←" together, we get $Tr_\beta(Tr_\gamma\dot{x}) \leftrightarrow Tr_\gamma(x)$ by ↔-Intro. And, since x was a fresh variable, we can derive:

$$\forall x(Tr_\beta(\underset{\cdot}{Tr}_\gamma\dot{x}) \leftrightarrow Tr_\gamma(x))$$

by ∀-Intro as desired.

- $\vdash_{PGA_{<\alpha}} \forall x(Sent_{<\gamma}(x) \rightarrow (Tr_\beta(\underset{\cdot}{Tr}_\gamma\dot{x}) \leftrightarrow Tr_\beta(x)))$ for $\gamma < \beta < \alpha$.

  Let x be a fresh variable for ∀-Intro. Assume $Sent_{<\gamma}(x)$ for a →-Intro. We now prove both directions of the biconditional $Tr_\beta(\underset{\cdot}{Tr}_\gamma\dot{x}) \leftrightarrow Tr_\beta(x)$.

  "→:" Assume $(*'')$ $Tr_\beta(\underset{\cdot}{Tr}_\gamma(\dot{x}))$ for →-Intro. From this, T_3^β, and PA, we can derive $Sent_{<\beta}(\underset{\cdot}{Tr}_\gamma\dot{x})$ and thus also $(**'')$ $Sent_{<\beta}(x)$. As before, we get $x \lhd \underset{\cdot}{Tr}_\gamma\dot{x}$ using $APD_T^{\beta,\gamma}$. Using (**") and G_{3a}^β, we can derive $Tr_\beta(x)$. Thus, we have $Tr_\beta(\underset{\cdot}{Tr}_\gamma\dot{x}) \rightarrow Tr_\beta(x)$ by →-Intro.

  "←:" Assume $Tr_\beta(x)$ for yet another →-Intro. From this and APU_T^β, we get $x \lhd Tr_\beta(x)$. Since we have assumed $Sent_{<\gamma}(x)$, we get $Tr_\gamma(x)$ from this and G_{3a}^γ. From this and APU_T^γ, we get $(*''')$ $x \lhd \underset{\cdot}{Tr}_\gamma(\dot{x})$. But now since $\gamma < \beta$, we can show in PA that $Sent_{<\beta}(\underset{\cdot}{Tr}_\gamma(\dot{x})$. But from this and $(*''')$, we get $Tr_\beta(\underset{\cdot}{Tr}_\gamma(\dot{x}))$. So, we have $Tr_\beta(x) \rightarrow Tr_\beta(\underset{\cdot}{Tr}_\gamma(\dot{x}))$ by →-Intro.

  Now putting both "→" and "←" together, we get $Tr_\beta(\underset{\cdot}{Tr}_\gamma(\dot{x})) \leftrightarrow Tr_\beta(x))$ by ↔-Intro and so $Sent_{<\gamma}(x) \rightarrow (Tr_\beta(\underset{\cdot}{Tr}_\gamma(\dot{x})) \leftrightarrow Tr_\beta(x))$ by →-Intro. Finally, since x was a fresh variable, we have:

$$\forall x(Sent_{<\gamma}(x) \rightarrow (Tr_\beta(\underset{\cdot}{Tr}_\gamma(\dot{x})) \leftrightarrow Tr_\beta(x))),$$

  as desired. □

□

This has the immediate consequence that for all ordinals $\beta < \alpha < \epsilon_0$, the theory $PGA_{<\alpha}$ proves the following typed version of the T-scheme for all languages $\mathcal{L}_{<\beta}$:

Lemma 2.8.5. *For all ordinals* $0 < \gamma \leq \beta < \alpha < \epsilon_0$ *and for all sentences* $\varphi \in \mathcal{L}_{<\gamma}$:

$$\vdash_{PGA_{<\alpha}} \forall t_1, \ldots, \forall t_n (Tr_\beta(\ulcorner \varphi(\dot{t_1}, \ldots, \dot{(t_n)}) \urcorner) \leftrightarrow \varphi(t_1, \ldots, t_n)).$$

Next, we will now show that for all ordinals $1 \leq \alpha < \epsilon_0$, the theory PGA_α is a proof-theoretically conservative extension of the theory PRT_α. But first, we need to introduce some more technical preliminaries: It is well-known that we can extend the technique of Gödel numbering to get terms for all ordinals below ϵ_0 [111, p. 17–42]. Let's denote the set of all ordinals below ϵ_0 by $On_{<\epsilon_0}$. We can adjust our coding function $\# : \mathcal{L} \to \mathbb{N}$ such that we injectively assign every ordinal $\alpha \in On_{<\epsilon_0}$ a unique code $\#\alpha \in \mathbb{N}$ that is different form all the codes $\#\sigma$ of the other expressions σ of $\mathcal{L}$. For all $\alpha \in On_{<\epsilon_0}$, we define the term $\ulcorner \alpha \urcorner$ to be $\overline{\#\alpha}$, i.e. our term for α is the numeral of the code $\#\alpha$ of α. Moreover, we extend the axioms of ordinary arithmetic to cover ordinal arithmetic up to ϵ_0. For simplicity, we'll use the same terminology for ordinal arithmetic and ordinary arithmetic. Thus, for example, we can now write $\ulcorner \alpha \urcorner \times \ulcorner \beta \urcorner$ in $\mathcal{L}$ to denote the product of ordinals $\alpha, \beta \in On_{<\epsilon_0}$. Moreover, we get:

$$\vdash_{PA} \ulcorner \alpha \urcorner \times \ulcorner \beta \urcorner = \ulcorner \gamma \urcorner \text{ iff } \alpha \times \beta = \gamma,$$

for all ordinals $\alpha, \beta, \gamma \in On_{<\epsilon_0}$. PA can represent the set of codes of ordinals below ϵ_0 and we'll use $\underset{\cdot}{On}_{<\epsilon_0}$ as a predicate for this. In particular, we get for all natural numbers $n \in \mathbb{N}$:

$$\vdash_{PA} \underset{\cdot}{On}_{<\epsilon_0}(\overline{n}) \text{ iff } n \in \#On_{<\epsilon_0} = \{\#\alpha \mid \alpha \in On_{<\epsilon_0}\}.$$

Finally, PA can prove the standard well-ordering $<$ of the ordinals below ϵ_0 and we'll use $\underset{\cdot}{<}$ to represent this ordering. We get for all ordinals $\alpha, \beta \in On_{<\epsilon_0}$:

$$\vdash_{PA} \ulcorner \alpha \urcorner \underset{\cdot}{<} \ulcorner \beta \urcorner \text{ iff } \alpha < \beta.$$

With these preliminaries in place,[47] we'll define a slightly non-standard notion of complexity for the formulas in $\mathcal{L}_{<\epsilon_0}$:

Definition 2.8.6 ('ω-complexity')**.** *For all ordinals* $1 \leq \alpha \leq \epsilon_0$, *we define the function* $|\ |_\omega : \mathcal{L}_{<\alpha} \to On_{<\epsilon_0}$ *that assigns to every formula* $\varphi \in \mathcal{L}_\alpha$ *its* ω*-complexity* $|\varphi|_\omega$ *recursively by saying that:*

(i) $|\varphi|_\omega = \begin{cases} \omega \times \alpha & \textit{if } \varphi = Tr_\alpha(t) \\ 0 & \textit{if } \varphi \textit{ atomic otherwise} \end{cases}$

(ii) $|\neg\varphi|_\omega = |\varphi|_\omega + 1$;

[47] Now we could define quantification over ordinals by saying that $\forall \gamma \varphi$ means $\forall x(\underset{\cdot}{On}_{<\epsilon_0}(x) \to \varphi)$ and work with the more general versions of the axioms mentioned before. But for reasons of perspicuity, we refrain from doing so.

(iii) $|\varphi \circ \psi|_\omega = lub(|\varphi|_\omega, |\psi|_\omega) + 1$, *for* $\circ = \wedge, \vee$; *and*

(iv) $|Qx\varphi|_\omega = |\varphi|_\omega + 1$, *for* $Q = \forall, \exists$.

Note that ω-complexity agrees with ordinary complexity on the formulas of $\mathcal{L}_{<1}$. Note furthermore that the function $x \mapsto \omega \times x$ is strictly monotonically increasing on the ordinals below ϵ_0:

Lemma 2.8.7. *For all* $\alpha, \beta \in On_{<\epsilon_0}$, *if* $\alpha < \beta$, *then* $\omega \times \alpha < \omega \times \beta$.

Note that as a consequence, we get that for all ordinals $0 < \alpha < \epsilon_0$, if $\varphi \in \mathcal{L}_{<\alpha}$, then $|\varphi|_\omega < |Tr_\alpha(\ulcorner\varphi\urcorner)|_\omega$. In other words, ω-complexity has a sort of "tracking property:" it can "track" the levels of Tarski's hierarchy. Moreover, we can represent ω-complexity in PA. More specifically, the function $c_\omega : \#\mathcal{L} \to \mathbb{N}$ that maps the code $\#\varphi$ of a formula $\varphi \in \mathcal{L}$ to its ω-complexity $|\varphi|_\omega$ is recursive and thus representable in PA. We represent c_ω by the unary function symbol $\underset{\cdot}{c_\omega}$. Thus, we get for all $\varphi \in \mathcal{L}_{<\epsilon_0}$ and all $\alpha \in On_{<\epsilon_0}$:

$$\vdash_{PA} \underset{\cdot}{c_\omega}(\ulcorner\varphi\urcorner) = \ulcorner\alpha\urcorner \text{ iff } |\varphi|_\omega = \alpha.$$

With this representation, we can prove the following provable version of the "tracking property" of ω-complexity:

Lemma 2.8.8. *For all ordinals* $0 < \beta \leq \alpha < \epsilon_0$*:*

$$\vdash_{PAG_\alpha} \forall x(Sent_{<\beta}(x) \to \underset{\cdot}{c_\omega}(x) \underset{\cdot}{<} \underset{\cdot}{c_\omega}(\underset{\cdot}{Tr_\beta}(\dot{x}))).$$[48]

Using ω-complexity, we'll obtain the main result of this section:

Theorem 2.8.9. *For all* $0 \leq \alpha < \epsilon_0$, *the theory* PGA_α *is a proof-theoretically conservative extension of the theory* PRT_α.

Proof. First, we define the translation function $\tau : \mathcal{L}_\alpha^{\lhd} \to \mathcal{L}_\alpha$ by saying that:

- $\tau(\varphi) = \begin{cases} Tr_\alpha(s) \wedge Tr_\alpha(t) \wedge \underset{\cdot}{c_\omega}(s) \underset{\cdot}{<} \underset{\cdot}{c_\omega}(t) & \text{if } \varphi = s \lhd t \\ \varphi & \text{if } \varphi \text{ atomic otherwise} \end{cases}$
- $\tau(\neg\varphi) = \neg\tau(\varphi)$;
- $\tau(\varphi \circ \psi) = \tau(\varphi) \circ \tau(\psi)$, for $\circ = \wedge, \vee$; and
- $\tau(Qx\varphi) = Qx(\tau(\varphi))$, for $Q = \forall, \exists$.

Then, we note that (a) for all $\varphi \in \mathcal{L}_\alpha$, $\tau(\varphi) = \varphi$. Next, we check that (b) for all $\varphi \in \mathcal{L}_\alpha^{\lhd}$, if $\vdash_{PGA_\alpha} \varphi$, then $\vdash_{PRT_\alpha} \tau(\varphi)$. The typed truth axioms of PGA_α are also axioms of PRT_α, so we only need to check the typed ground axioms and the typed upward and downward directed axioms. Here we only go through a few cases to illustrate the idea:

[48] We don't give the detailed proof here, but it essentially proceeds by using induction on ordinals below ϵ_0 in PA.

- In the case of the axiom G_4^α, we get:

  $\tau(G_4^\alpha) = \forall x \forall y((Tr_\alpha(x) \wedge Tr_\alpha(y) \wedge c_{\underset{\cdot}{\omega}}(x) \underset{\cdot}{<} c_{\underset{\cdot}{\omega}}(y)) \to Sent_{<\alpha}(x) \wedge Sent_{<\alpha}(y))$

  This is provable (almost) immediately from the typed truth axiom T_3^α of PRT_α:

  $$\forall x(Tr_\alpha(x) \to Sent_{<\alpha}(x)).$$

- Finally, consider the axioms (APU_T^β):

  $$\forall x(Tr_\beta(x) \to x \lhd \underset{\cdot}{Tr}_\beta(\dot{x})),$$

  where $\beta < \alpha$. We get:

  $\tau(APU_T^\beta) = \forall x(Tr_\beta(x) \to Tr_\alpha(x) \wedge Tr_\alpha(\underset{\cdot}{Tr}_\beta(\dot{x})) \wedge c_{\underset{\cdot}{\omega}}(x) \underset{\cdot}{<} c_{\underset{\cdot}{\omega}}(\underset{\cdot}{Tr}_\beta(\dot{x}))).$

  Now let x be a fresh variable for $\forall$-Intro and assume $Tr_\beta(x)$ for a $\to$-Intro. Using the axiom T_3^β of PRT_α, we can infer that $Sent_{<\beta}(x)$. Moreover, since $\beta < \alpha$ by assumption, we can infer that $Tr_\alpha(\underset{\cdot}{Tr}_\beta(\dot{x}))$ and $Tr_\beta(x)$ using the axiom RP_{12}^α of PRT_α. Finally, by Lemma 2.8.8, we get $Sent_{<\beta}(x) \to c_{\underset{\cdot}{\omega}}(x) \underset{\cdot}{<} \underset{\cdot}{Tr}_\beta(\dot{x})$. Since we know already that $Sent_{<\beta}(x)$, we get the final piece $c_{\underset{\cdot}{\omega}}(x) \underset{\cdot}{<} c_{\underset{\cdot}{\omega}} \underset{\cdot}{Tr}_\beta(\dot{x})$. Putting all of this together, by $\to$-Intro, we have

  $$Tr_\beta(x) \to Tr_\alpha(x) \wedge Tr_\alpha(\underset{\cdot}{Tr}_\beta(\dot{x})) \wedge c_{\underset{\cdot}{\omega}}(x) \underset{\cdot}{<} c_{\underset{\cdot}{\omega}}(\underset{\cdot}{Tr}_\beta(\dot{x})),$$

  and since x was a fresh variable, by $\forall$-Intro, we get the desired theorem.

Putting (a) and (b) together, the claim follows.

□ □

The theorem has the following immediate consequence:[49]

Corollary 2.8.10. *For all $0 \leq \alpha < \epsilon_0$, the theory PGA_α is consistent.*

The proof of Theorem 2.8.9 essentially works because of the "tracking property" of ω-complexity. The idea of the proof is the same as in the proof of the corresponding result in the first part of this chapter, but the translation we used there would not have worked. Sentences of the form $Tr_\beta(\ulcorner \varphi \urcorner)$ involving the truth predicate all have a classical complexity of zero, while the sentence

[49] We could also use the theorem to determine the proof theoretic strength of PGA_α, but there is a small "hickup:" the version of PRT_α that we discussed here is not exactly the one that is usually discussed in the literature. As mentioned in Footnote 46, PRT_α is usually formulated using quantification over ordinals, which we avoided here for reasons of perspicuity. The version of PRT_α with axioms quantifying over ordinals proves the same arithmetical theorems as the theory RA_α of *ramified analysis up to (and including) α*. For a proof of this result, see [35]. We suspect that the proof theoretic strength of our version of PRT_α is very close to this, although we're not going to prove anything to this effect.

φ may have arbitrary complexity. Thus, we would not be able to derive the translations of the (typed versions of the) Aristotelian principles under the translation from the previous chapter. The trick was to use ω-complexity in the translation—this is what allowed us to prove the result. The technique of the proof works for all ordinals $\alpha < \epsilon_0$, since PA can prove the well-ordering of these ordinals, which is required for the proof. The theory $PGA_{<\epsilon_0}$ is the first theory where our proof doesn't work anymore, because in this theory we don't have a "highest" truth predicate as required for the definition of τ. But we can extend our result to this theory using a simple *compactness argument*:

Corollary 2.8.11. *The theory $PGA_{<\epsilon_0}$ is a proof-theoretically conservative extension of the theory $PRT_{<\epsilon_0}$.*

Proof. Assume that there is a sentence $\varphi \in \mathcal{L}_{<\epsilon_0}$ such that $\vdash_{PGA_{<\epsilon_0}} \varphi$, but $\nvdash_{PRT_{<\epsilon_0}} \varphi$. Since proofs are finite objects, there can only be finitely many occurrences of different truth predicates $Tr_{\beta_1}, \ldots, Tr_{\beta_n}$, for $0 < \beta_1 < \ldots < \beta_n < \epsilon_0$, in the proof. But then the proof of φ, is also a proof in PGA_{β_n} and $\varphi \in \mathcal{L}_{\beta_n}$. Now by Theorem 2.8.9, PGA_{β_n} is conservative over PRT_{β_n}. This means that $\vdash_{PRT_{\beta_n}} \varphi$ and thus also $\vdash_{PRT_{<\epsilon_0}} \varphi$. Contradiction! Thus, there is no such φ and the claim holds. $\square$ $\square$

We get immediately:

Corollary 2.8.12. *The theory $PGA_{<\epsilon_0}$ is consistent.*

The theory $PGA_{<\epsilon_0}$ is a natural stopping point for the methods we've developed in this chapter.[50]

We have shown the consistency of our theories $PGA_{<\alpha}$, where $1 \leq \alpha \leq \epsilon_0$, by proof theoretic means. But for reasons of perspicuity, it would also be good to have an idea what models for these theories look like. In the rest of this section, we will show how to extend the construction from the previous chapter to obtain models for $PGA_{<\alpha}$, where $1 \leq \alpha \leq \epsilon_0$.

As in the case of PT, there is a standard model of $PRT_{<\alpha}$, for $1 \leq \alpha \leq \epsilon_0$. A model for the language $\mathcal{L}_{<\alpha}$ is a structure of the form $(\mathbf{N}, (S_\beta)_{\beta<\alpha})$, where for $\beta < \alpha$, the set S_β interprets the truth predicate $Tr_\beta \in \mathcal{L}_{<\alpha}$. For $1 \leq \alpha \leq \epsilon_0$, we define the sets $(\mathbf{S}_\beta)_{\beta<\alpha}$ by the following (transfinite) recursion:

- $\mathbf{S}_1 = \{\#\varphi \mid \varphi \in \mathcal{L}_{<1}, \mathbf{N} \vDash \varphi\}$;
- $\mathbf{S}_{\alpha+1} = \mathbf{S}_\alpha \cup \{\#\varphi \mid \varphi \in \mathcal{L}_{<\alpha}, (\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha}) \vDash \varphi\}$
- $\mathbf{S}_\alpha = \bigcup_{\beta<\alpha} \mathbf{S}_\beta$, if α is a limit ordinal.

[50] The situation is quite similar to the corresponding theories of truth. For a discussion of the natural stopping point see [58, p. 322-29].

Then we get, for all $1 \leq \alpha \leq \epsilon_0$, that $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha}) \vDash PRT_{<\alpha}$. For $1 \leq \alpha \leq \epsilon_0$, the model $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ is the *standard model* of $PRT_{<\alpha}$—it models Tarski's hierarchy of truths.

We now extend our definition of grounding-trees from the previous chapter to *grounding-trees over the standard model of* $PRT_{<\alpha}$:

Definition 2.8.13. *Let $1 \leq \alpha \leq \epsilon_0$ and let $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ be the standard model of $PRT_{<\alpha}$. We define the grounding-trees over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ by the following clauses for all formulas $\varphi \in \mathcal{L}_{<\alpha}$:*

(i) *$\#\varphi \in \bigcup_{\beta<\alpha} \mathbf{S}_\beta$, then $\#\varphi$ is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\varphi$ as its root;*

(ii) *if $\begin{array}{c}\#\varphi\\ \widehat{\mathcal{T}}\end{array}$ is a grounding-tree $\mathcal{T}$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\varphi$ as its root, then*

$$\begin{array}{c}\#\neg\neg\varphi\\ |\\ \#\varphi\\ \widehat{\mathcal{T}}\end{array}$$

is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\neg\neg\varphi$ as its root;

(iii) *if $\begin{array}{c}\#\varphi\\ \widehat{\mathcal{T}}\end{array}$ is a grounding-tree $\mathcal{T}$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\varphi$ as its root, then*

$$\begin{array}{c}\#(\varphi \vee \psi)\\ |\\ \#\varphi\\ \widehat{\mathcal{T}}\end{array}$$

is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#(\varphi \vee \psi)$ as its root;

(iv) *if $\begin{array}{c}\#\psi\\ \widehat{\mathcal{T}}\end{array}$ is a grounding-tree $\mathcal{T}$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\psi$ as its root, then*

$$\begin{array}{c}\#(\varphi \vee \psi)\\ |\\ \#\psi\\ \widehat{\mathcal{T}}\end{array}$$

is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#(\varphi \vee \psi)$ as its root;

(v) *if $\begin{array}{c}\#\varphi\\ \widehat{\mathcal{T}_1}\end{array}$, $\begin{array}{c}\#\psi\\ \widehat{\mathcal{T}_2}\end{array}$ are grounding-trees $\mathcal{T}_1, \mathcal{T}_2$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\varphi, \#\psi$*

as their roots respectively, then $\begin{array}{c}\#(\varphi\wedge\psi)\\ \diagup\quad\diagdown\\ \#\varphi\qquad\#\psi\\ \widehat{\mathcal{T}_1}\qquad\widehat{\mathcal{T}_2}\end{array}$ is a grounding-tree over $(\mathbf{N},(\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#(\varphi\wedge\psi)$ as its root;*

(vi) if $\begin{array}{c}\#\varphi(t)\\ \widehat{\mathcal{T}}\end{array}$ is a grounding-tree $\mathcal{T}$ over $(\mathbf{N},(\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\varphi(t)$ as its root, then $\begin{array}{c}\#\exists x\varphi(x)\\ |\\ \#\varphi(t)\\ \widehat{\mathcal{T}}\end{array}$ is a grounding-tree over $(\mathbf{N},(\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\exists x\varphi(x)$ as its root;

(vii) if $\begin{array}{c}\#\varphi(t_1)\\ \widehat{\mathcal{T}_1}\end{array}$, $\begin{array}{c}\#\varphi(t_2)\\ \widehat{\mathcal{T}_2}\end{array}$, $\ldots$ are grounding-trees $\mathcal{T}_1,\mathcal{T}_2,\ldots$ over $(\mathbf{N},(\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\varphi(t_1),\#\varphi(t_2),\ldots$ as their roots respectively, where $t_1,t_2,\ldots$ are all and only the terms of $\mathcal{L}_{PA}$, then $\begin{array}{c}\#\forall x\varphi(x)\\ \diagup\quad|\quad\diagdown\\ \#\varphi(t_1)\quad\#\varphi(t_2)\quad\ldots\\ \widehat{\mathcal{T}_1}\qquad\widehat{\mathcal{T}_2}\qquad\end{array}$ is a grounding-tree over $(\mathbf{N},(\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\forall x\varphi(x)$ as its root;

(viii) if $\begin{array}{c}\#\neg\varphi\\ \widehat{\mathcal{T}}\end{array}$ is a grounding-tree $\mathcal{T}$ over $(\mathbf{N},(\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\neg\varphi$ as its root, then $\begin{array}{c}\#\neg(\varphi\wedge\psi)\\ |\\ \#\neg\varphi\\ \widehat{\mathcal{T}}\end{array}$ is a grounding-tree over $(\mathbf{N},(\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\neg(\varphi\wedge\psi)$ as its root;

(ix) if $\begin{array}{c}\#\neg\psi\\ \widehat{\mathcal{T}}\end{array}$ is a grounding-tree $\mathcal{T}$ over $(\mathbf{N},(\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\neg\psi$ as its root, then $\begin{array}{c}\#\neg(\varphi\wedge\psi)\\ |\\ \#\neg\psi\\ \widehat{\mathcal{T}}\end{array}$ is a grounding-tree over $(\mathbf{N},(\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\neg(\varphi\wedge\psi)$ as its root;

(x) *if* $\underset{\widehat{\ \mathcal{T}_1\ }}{\#\neg\varphi}$, $\underset{\widehat{\ \mathcal{T}_2\ }}{\#\neg\psi}$ *are grounding-trees* $\mathcal{T}_1, \mathcal{T}_2$ *over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#\neg\varphi, \#\neg\psi$ *as their roots respectively, then* $\begin{array}{c} \#\neg(\varphi \vee \psi) \\ \diagup \quad \diagdown \\ \underset{\widehat{\ \mathcal{T}_1\ }}{\#\neg\varphi} \quad \underset{\widehat{\ \mathcal{T}_2\ }}{\#\neg\psi} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#\neg(\varphi \vee \psi)$ *as its root;*

(xi) *if* $\underset{\widehat{\ \mathcal{T}\ }}{\#\neg\varphi(t)}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#\neg\varphi(t)$ *as its root, then* $\begin{array}{c} \#\neg\forall x\varphi(x) \\ | \\ \underset{\widehat{\ \mathcal{T}\ }}{\#\neg\varphi(t)} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#\neg\forall x\varphi(x)$ *as its root;*

(xii) *if* $\underset{\widehat{\ \mathcal{T}_1\ }}{\#\neg\varphi(t_1)}$, $\underset{\widehat{\ \mathcal{T}_2\ }}{\#\neg\varphi(t_2)}$, $\ldots$ *are grounding-trees* $\mathcal{T}_1, \mathcal{T}_2, \ldots$ *over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#\neg\varphi(t_1), \#\neg\varphi(t_2), \ldots$ *as their roots respectively, where* $t_1, t_2, \ldots$ *are all and only the terms of* $\mathcal{L}_{PA}$, *then* $\begin{array}{c} \#\neg\exists x\varphi(x) \\ \diagup \quad | \quad \diagdown \\ \underset{\widehat{\ \mathcal{T}_1\ }}{\#\neg\varphi(t_1)} \quad \underset{\widehat{\ \mathcal{T}_2\ }}{\#\neg\varphi(t_2)} \quad \ldots \end{array}$ *is a grounding-tree over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#\forall x\varphi(x)$ *as its root;*

(xiii) *if* $\underset{\widehat{\ \mathcal{T}\ }}{\#\varphi}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#\varphi$ *as its root and* $\#\varphi \in \mathbf{S}_\beta$, *for* $\beta < \alpha$, *then* $\begin{array}{c} \#Tr_\beta(\ulcorner\varphi\urcorner) \\ | \\ \underset{\widehat{\ \mathcal{T}\ }}{\#\varphi} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#Tr_\beta(\ulcorner\varphi\urcorner)$ *as its root;*

(xiv) *if* $\underset{\widehat{\ \mathcal{T}\ }}{\#\neg\varphi}$ *is a grounding-tree* $\mathcal{T}$ *over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#\neg\varphi$ *as its root*

and $\#\varphi \in \mathbf{S}_\beta$*, for* $\beta < \alpha$*, then* $\begin{array}{c} \#\neg Tr_\beta(\ulcorner \varphi \urcorner) \\ | \\ \#\varphi \\ \widehat{\mathcal{T}} \end{array}$ *is a grounding-tree over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *with* $\#\neg Tr_\beta(\ulcorner \varphi \urcorner)$ *as its root;*

(xv) nothing else is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$.

Now, in contrast to grounding-trees over $(\mathbf{N}, \mathbf{S})$, grounding-trees over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ can have an *infinite* height:

Definition 2.8.14. *We define the height* $h(\mathcal{T})$ *of a grounding tree over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *by saying that:*

(i) all grounding-trees over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *of the form* $\#\varphi \in \mathbf{S}$ *have height one;*

(ii) if $\mathcal{T}$ *is a grounding-tree over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *that is constructed from grounding-trees* $\mathcal{T}_1, \mathcal{T}_2, \ldots$ *over* $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$*, then the height of* $\mathcal{T}$ *is one plus the least upper bound of the heights of* $\mathcal{T}_1, \mathcal{T}_2, \ldots$*:*

$$h(\mathcal{T}) = lub\{h(\mathcal{T}_1), h(\mathcal{T}_2), \ldots\} + 1,$$

where lub is the operation of taking the least upper bound.

We call a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ *degenerate iff it is of height one.*

To see that there are grounding-trees of infinite height, leet $DN_{\overline{0}=\overline{0}}(x)$ represent the property of being an instance of $\overline{0} = \overline{0}$ preceded by an even number of negations. Then it is easily checked that for all φ such that $DN_{\overline{0}=\overline{0}}(\#\varphi)$, there is a grounding-tree of the form

$$\begin{array}{c} DN_{\overline{0}=\overline{0}}(\ulcorner \neg \ldots \neg(\overline{0} = \overline{0}) \urcorner) \to \#Tr_1(\ulcorner \neg \ldots \neg(\overline{0} = \overline{0}) \urcorner) \\ | \\ \#Tr_1(\ulcorner \neg \ldots \neg(\overline{0} = \overline{0}) \urcorner) \\ | \\ \vdots \\ | \\ \#\overline{0} = \overline{0} \end{array}$$

which has height $\frac{n}{2} + 3$, where n is the number of negations in φ. A consequence of this is that the least upper bound of the heights of $\mathcal{T}_1, \mathcal{T}_2, \ldots$ in the grounding-tree

$$\#\forall x(DN_{\overline{0}=\overline{0}}(x) \rightarrow \#Tr_1(x))$$

$$\underbrace{\#DN_{\overline{0}=\overline{0}}(t_1) \rightarrow \#Tr_1(t_1)}_{\widehat{\mathcal{T}_1}} \quad \underbrace{\#DN_{\overline{0}=\overline{0}}(t_2) \rightarrow \#Tr_1(t_2)}_{\widehat{\mathcal{T}_2}} \quad \ldots$$

is at least ω and thus the height of this tree is at least $\omega + 1$.[51]

Now, an important consequence of this observation is that we can't use ordinary induction on the height of trees to prove claims about all grounding-trees. We need to use *transfinite* induction on the height of the grounding-trees. This doesn't add any further complications, but to be explicit let's state the form of the principle that we're going to use. Consider a property of grounding-trees. Then, if we can show that any degenerate grounding-tree has the property and we can show that if we can show that assuming that all trees of a height smaller than a given tree have the property, then the tree itself has the property, it follows that all grounding-trees have the property. Note that in this form of the principle, the induction step also includes *limit cases*, where we consider a tree of the height of a limit ordinal and need to show that the tree has the property in question, given that all trees of a lower height have the property.

Analogously to the case of grounding-trees over $(\mathbf{N}, \mathbf{S})$, we can now show that grounding-trees over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ are: (i) *rooted graphs* over $\bigcup_{\beta<\alpha} \mathbf{S}_\beta$; (ii) indeed *rooted trees* over $\bigcup_{\beta<\alpha} \mathbf{S}_\beta$, i.e. they don't contain any cycles; and finally, (iii) transitive.

Lemma 2.8.15. *Let $1 \leq \alpha \leq \epsilon_0$, $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ be the standard model of $PRT_{<\alpha}$, and let $\mathcal{T}$ be a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$. Then for all formulas $\varphi \in \mathcal{L}_{<\alpha}$, if $\#\varphi$ is a vertex in $\mathcal{T}$, then $\#\varphi \in \bigcup_{\beta<\alpha} \mathbf{S}_\beta$.*

Proof. The new cases for clauses (xiii) and (xiv) follow by the fact that $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ is a model of $PRT_{<\alpha}$. □ □

Remember the notion of a code of a formula occurring below another code in a grounding-tree over $(\mathbf{N}, \mathbf{S})$. We now adapt this notion to grounding-trees over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ by recursively saying that, for all $1 \leq \alpha \leq \epsilon_0$, no code of any formula occurs below the code of any other formula in a degenerate grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$, and if $\mathcal{T}$ is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ that was constructed from grounding-trees $\mathcal{T}_1, \mathcal{T}_2, \ldots$ over

[51] In fact, by a slightly more complicated argument we can show that the height of this tree is *exactly* $\omega + 1$.

$(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ according to the rules (ii–xvi) of Definition 2.8.13, then all occurrences of all formulas in $\mathcal{T}_1, \mathcal{T}_2, \ldots$ occur below the root of $\mathcal{T}$ in $\mathcal{T}$.

Then we can show:

Lemma 2.8.16. *Let $1 \leq \alpha \leq \epsilon_0$ and let $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ be the standard model of $PRT_{<\alpha}$. If $\mathcal{T}$ is a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\varphi$ as its root, for some formula $\varphi \in \mathcal{L}_{<\alpha}$. Then, all formulas $\psi \in \mathcal{L}_{<\alpha}$ whose code $\#\psi$ occurs below $\#\varphi$ in $\mathcal{T}$ have a lower ω-complexity than φ.*

Lemma 2.8.17. *Let $1 \leq \alpha \leq \epsilon_0$, $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ be the standard model of $PRT_{<\alpha}$, and let $\mathcal{T}$ be a grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$. Then between any two nodes $\#\varphi$ and $\#\psi$ in $\mathcal{T}$, for formulas $\varphi, \psi \in \mathcal{L}_{<\alpha}$, there is exactly one path.*

Lemma 2.8.18. *Let $1 \leq \alpha \leq \epsilon_0$ and let $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ be the standard model of $PRT_{<\alpha}$. If there is a grounding-tree $\mathcal{T}_1$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\psi$ as its root and $\#\varphi_1, \#\varphi_2, \ldots$ as its leaves and there is grounding-tree $\mathcal{T}_2$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\psi, \#\psi_1, \#\psi_2, \ldots$ as its leaves and $\#\theta$ as its root, then there is a grounding-tree $\mathcal{T}_3$ over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with $\#\varphi_1, \#\varphi_2, \ldots, \#\psi_1, \#\psi_2, \ldots$ as its leaves and $\#\theta$ as its root.*

Finally, we define the standard model of $PGA_{<\alpha}$ by saying that:

Definition 2.8.19. *Let $1 \leq \alpha \leq \epsilon_0$ and let $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ be the standard model of $PRT_{<\alpha}$. We define the relation $\mathbf{R} \subseteq \mathbb{N}^2$ by saying that for all $n, m \in \mathbb{N}$, $\mathbf{R}(m, n)$ iff there is a non-degenerate grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ with n as a leaf and m as its root.*

Putting Lemmas 2.8.15, 2.8.17, and 2.8.18 together, we obtain:

Theorem 2.8.20. *$(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha}, \mathbf{R})$ is a model of $PGA_{<\alpha}$, for $1 \leq \alpha \leq \epsilon_0$, i.e. $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha}, \mathbf{R}) \vDash PGA_{<\alpha}$.*

Proof. By Lemmas 2.8.15, 2.8.17, and 2.8.18, grounding-trees over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ behave appropriately and satisfy the basic ground axioms. Since $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$ is a model of $PRT_{<\alpha}$, the typed truth axioms are satisfied. The only new cases are the axioms for the typed Aristotelian principles. Here we only show that APU_T^γ, for $\gamma < \alpha$ holds:

- $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha}, \mathbf{R}) \vDash \forall x(Tr_\gamma(x) \to x \lhd \underset{\cdot}{Tr_\gamma}\dot{x})$ for all $\gamma < \alpha$.

  Let σ be a variable assignment over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha}, \mathbf{R})$ and σ' some x-variant of σ. Assume that $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha}, \mathbf{R}) \vDash_{\sigma'} Tr_\gamma(x)$. This means that $\sigma'(x) \in \mathbf{S}_\gamma$. Since $\mathbf{S}_\gamma = \{\#\varphi \mid \varphi \in \mathcal{L}_{<\gamma}, (\mathbf{N}, (\mathbf{S}_\delta)_{\delta<\gamma}) \vDash \varphi\}$, we know that $\sigma'(x) = \#\varphi$, for some formula $\varphi \in \mathcal{L}_{<\gamma}$. Now, $\#\varphi$ is a degenerate grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$. But then, by clause (xiii) of Definition 2.8.13, $\begin{array}{c}\#Tr_\gamma(\ulcorner\varphi\urcorner)\\ \mid \\ \#\varphi\end{array}$ is a non-degenerate grounding-tree over $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha})$. Moreover, the root of this tree is $\#Tr_\gamma(\ulcorner\varphi\urcorner)$ and its only leaf is $\#\varphi$. Now consider $\sigma'(\dot{Tr_\gamma}\dot{x})$. Since we know that $\sigma'(x) = \#\varphi$ and $\underset{\cdot}{Tr_\gamma}$ expresses the function that maps codes of formulas to the code of $\underset{\cdot}{Tr_\gamma}$ applied to the formula, we know that $\sigma'(\underset{\cdot}{Tr_\gamma}\ulcorner\dot{\varphi}\urcorner) = \#Tr_\gamma(\ulcorner\varphi\urcorner)$. Thus, $\mathbf{R}(\sigma'(x), \sigma'(\underset{\cdot}{Tr_\gamma}\dot{x}))$ meaning $\vDash_{\sigma'} x \lhd \underset{\cdot}{Tr_\gamma}\dot{x}$. And since σ was arbitrary, we get $(\mathbf{N}, (\mathbf{S}_\beta)_{\beta<\alpha}, \mathbf{R}) \vDash \forall x(Tr_\gamma(x) \to x \lhd \underset{\cdot}{Tr_\gamma}\dot{x})$, as desired.

We can show analogously that the other axioms hold.

□ □

2.9 Paradoxes of Self-Referential Ground

In the setting of PG (as well as $PGA_{<\alpha}$ for $0 < \alpha < \epsilon_0$), we can accommodate new n-ary predicates R by stipulating:

$$\forall t_1, \ldots, \forall t_n(Tr(\underset{\cdot}{R}(t_1, \ldots, t_n)) \leftrightarrow R(t_1^\circ, \ldots, t_n^\circ)) \text{ and}$$

$$\forall t_1, \ldots, \forall t_n(Tr(\underset{\cdot}{\neg}\underset{\cdot}{R}(t_1, \ldots, t_n)) \leftrightarrow \neg R(t_1^\circ, \ldots, t_n^\circ)).$$

If we furthermore add a theory for the new predicate R to our background theory, the resulting theory will be a theory of partial ground over the hierarchy of truths of that extended background theory. For example, we could formulate theories of partial ground over the truths of analysis, of mereology, or even of set-theory.[52] The resulting theory will then, of course, also prove all the instances of the T-scheme $Tr(\ulcorner\varphi\urcorner) \leftrightarrow \varphi$ over sentences of the new language $\mathcal{L} \cup \{R\}$. But this approach has limits. Of course, we can't let R be our unary truth predicate Tr itself. It is well-known that if we affirm:

$$\forall t(Tr(\underset{\cdot}{Tr}(t)) \leftrightarrow Tr(t^\circ)) \text{ and}$$

$$\forall t(Tr(\underset{\cdot}{\neg}\underset{\cdot}{Tr}(t)) \leftrightarrow \neg Tr(t^\circ)),$$

our theory will fall prey to the liar paradox and its ilk. It might be somewhat surprising to learn, however, that we also can't affirm:

$$\forall s \forall t(Tr(s \underset{\cdot}{\lhd} t) \leftrightarrow s^\circ \lhd t^\circ) \text{ and}$$

$$\forall s \forall t(Tr(s \underset{\cdot}{\not\lhd} t) \leftrightarrow s^\circ \not\lhd t^\circ)$$

in the context of predicational theories of ground. This will be the main result of this section.[53] We'll call the resulting problem the *new puzzle of ground*.

[52]The truths of new atomic sentences will, of course, not have any provable grounds in the theory.

[53]Here $s \not\lhd t$, for terms s and t, is an abbreviation of $\neg(s \lhd t)$, analogous to the case of $s \neq t$. Similarly, the notation $s \underset{\cdot}{\not\lhd} t$ is an abbreviation for the complex function term $\underset{\cdot}{\neg}(s \underset{\cdot}{\lhd} t)$ for terms s and t.

Let's move to a setting where we can affirm applications of the ground predicate to sentences involving the ground predicate. For this purpose, we have to make a couple of adjustments to our theory setup. In the following, let $\mathcal{L}_\lhd$ be the language $\mathcal{L} \cup \{\lhd\}$. Now first, we have to assume that we work in the context of a proper coding for $\mathcal{L}_\lhd$. In particular, we now assume that we have a name $\ulcorner\varphi\urcorner$ for every sentence $\varphi \in \mathcal{L}_\lhd$. We assume that we have a function symbol $\underset{\cdot}{\lhd}$ such that:

$$\vdash_{PA} s \underset{\cdot}{\lhd} t = \ulcorner s \lhd t \urcorner,$$

for all terms s and t. And we let $Sent_\lhd$ abbreviate the formula that allows us to represent the set (of codes of) sentences of $\mathcal{L}_\lhd$. In particular, we assume that for all $n \in \mathbb{N}$:

$$\vdash_{PA} Sent_\lhd(\overline{n}) \text{ iff } n \in \#\mathcal{L}_\lhd \text{ and}$$
$$\vdash_{PA} \neg Sent_\lhd(\overline{n}) \text{ iff } n \notin \#\mathcal{L}_\lhd.$$

Second, we need to adjust the axiom T_3 to:

$$(\mathrm{T}_3^\lhd) \quad \forall x(Tr(x) \to Sent_\lhd(x)),$$

which allows sentences involving the ground predicate to occur in the context of the truth predicate (and thus in the context of the ground predicate). We arrive at a modified predicational ground:

Definition 2.9.1. *The predicational theory PUG of untyped ground consists of the axioms of PG without the axiom T_3, plus the axiom $T_3^\lhd$ and the axioms:*

$$(T_\lhd^1) \quad \forall s \forall t(Tr(s \underset{\cdot}{\lhd} t) \leftrightarrow s^\circ \lhd t^\circ) \text{ and}$$

$$(T_\lhd^2) \quad \forall s \forall t(Tr(s \underset{\cdot}{\ntriangleleft} t) \leftrightarrow s^\circ \ntriangleleft t^\circ).$$

We will now show that PUG is inconsistent. To see this, first note that we can then show in the same way as in the case of PG that PUG proves the uniform T-scheme for sentences involving the ground predicate:

Lemma 2.9.2. *For all sentences $\varphi \in \mathcal{L}_\lhd$,*

$$\vdash_{PUG} \forall t_1, \ldots, \forall t_n(Tr(\varphi(\underset{\cdot}{t_1}, \ldots, \underset{\cdot}{t_n})) \leftrightarrow \varphi(t_1^\circ, \ldots, t_n^\circ)).$$

Proof. By induction on the positive complexity of formulas. The new axioms $T_\lhd^{1/2}$ take care of the new base-case. □ □

Note that this lemma doesn't entail yet that PUG is inconsistent: the truth predicate Tr is not in the language $\mathcal{L}_\lhd$ and thus Lemma 2.9.2 doesn't entail that we're applying the truth predicate to sentences involving the same truth predicate. To see that PUG is inconsistent, we need to do some more work. Since we work in the context of a coding for $\mathcal{L}_\lhd$, we can get the diagonal lemma for the language:

Lemma 2.9.3. *For all formulas $\varphi \in \mathcal{L}_{\lhd}$ with exactly one free variable, there is a formula $\delta \in \mathcal{L}_{\lhd}$ such that*

$$\vdash_{PA} \delta \leftrightarrow \varphi(\ulcorner \delta \urcorner).$$

With these two lemmas in place, we can show our main result:
Theorem 2.9.4. *PUG is inconsistent.*

Proof. Let $\varphi(x)$ be the formula $\neg(x \lhd \underset{\cdot}{\neg}\underset{\cdot}{\neg} x)$. By the diagonal lemma for $\mathcal{L}_{\lhd}$, we know that there is a formula $\delta \in \mathcal{L}_{\lhd}$ such that:

$$\vdash_{PA} \delta \leftrightarrow \neg(\ulcorner \delta \urcorner \lhd \ulcorner \neg\neg\delta \urcorner).$$

It is easily checked that we can now both prove δ and $\neg\delta$. $\square$

$\square$

We are left with yet another puzzle of ground: by letting the truth predicate (and thus the ground predicate) apply to truths involving the ground predicate, we made our intuitively plausible theory of ground inconsistent. But intuitively, we want to be able to talk about the truth of sentences involving the ground predicate. So what went wrong?

First, note that the new puzzle is different from Fine's puzzle of ground. Fine's puzzle consists in the fact that different intuitively plausible principles for partial ground and truth entails that the truth of some sentences partially ground themselves—in contradiction to the irreflexivity of partial ground. Our puzzle, in contrast, consists in the fact that letting the truth predicate apply to the ground predicate makes our previously consistent principles of ground inconsistent.

Moreover, note that the use of double negation (and of the corresponding upward directed ground axiom) in the proof of Theorem 2.9.4 is dispensable. We could equally well have applied the diagonal lemma to the formula $\neg(x \lhd x \underset{\cdot}{\vee} x)$ or $\neg(x \lhd x \underset{\cdot}{\wedge} x)$ or ... and we would still have gotten the same inconsistency result (using the corresponding upward directed ground axioms for these connectives). The point is that our paradox is not a paradox of double negation, or disjunction, or conjunction, or the like—it has another source.

We argue that the paradox is a paradox of *self-reference* in the context of partial ground. In support of this claim, first note that all the different sentences that we could use for the proof of Theorem 2.9.4 have in common that they (provably) "say of themselves" that they violate certain principles of partial ground, and as a consequence they are inconsistent over our theory of untyped ground. Thus, the new puzzle of ground bears a strong resemblance to the classic semantic paradoxes. Indeed, we can make this analogy even more explicit in the following proposition. Let's define the predicate $Tr_0^=$ by the following explicit definition:

$$\forall x(Tr_0^=(x) \leftrightarrow_{def} \exists s,t(x = (s \dot{=} t) \wedge s^\circ = t^\circ) \vee \exists s,t(x = (s \dot{\neq} t) \wedge s^\circ \neq t^\circ))$$

And let's define the predicate $Tr_0^\lhd$ by the following explicit definition:

$$\forall x(Tr_0^\lhd(x) \leftrightarrow_{def} \exists s,t(x = (s \dot{\lhd} t) \wedge s^\circ \lhd t^\circ) \vee \exists s,t(x = (s \dot{\not\lhd} t) \wedge s^\circ \not\lhd t^\circ))$$

So, intuitively $Tr_0^=$ and $Tr_0^\lhd$ are truth predicates for the literals of $\mathcal{L}_\lhd$. We can then show the following proposition:

Proposition 2.9.5. *PUG proves that the formula* $Tr_0^=(x) \vee Tr_0^\lhd(x) \vee \exists y(y \lhd x)$ *satisfies the T-scheme for the formulas of* $\mathcal{L}_\lhd$*, i.e. for all* $\varphi \in \mathcal{L}_\lhd$*:*

$$\vdash_{PUG} Tr_0^=(\ulcorner\varphi\urcorner) \vee Tr_0^\lhd(\ulcorner\varphi\urcorner) \vee \exists y(y \lhd \ulcorner\varphi\urcorner) \leftrightarrow \varphi.$$

Proof. By induction on the positive complexity of φ for both directions of the biconditional. The base cases for φ being $s = t$ or $s \neq t$, for terms s and t, are covered by the basic truth axioms T_1 and T_2 and the definition of $Tr_0^=$. Similarly, the base cases for φ being $s \lhd t$ or $s \not\lhd t$, for terms s and t, are covered by the axioms $T_\lhd^1$ and $T_\lhd^2$ and the definition of $Tr_0^\lhd$. The remaining cases are can be dealt with using Lemma 2.9.2. □ □

In other words, in *PUG* we can define a truth predicate for $\mathcal{L}_\lhd$ that satisfies the T-scheme for $\mathcal{L}_\lhd$. It follows by Tarski's *theorem of the undefinability of truth* that *PUG* is inconsistent [132]. Our Theorem 2.9.4 is only a special case of this more general fact, as it were. This is the precise sense in which the ground predicate "behaves too much like a truth predicate"—in other words: the new puzzle of ground is at heart a paradox of self-reference.

2.9.1 Conclusions

How should we respond to the paradox of self-referential ground? Three natural ways in which we could try to block the inconsistency theorem suggest themselves: First, we could try to rule out self-referential sentences of ground like the one used in the proof of the inconsistency theorem. Second, we could try to restrict the principles of partial ground used in the proof of the inconsistency theorem. And third, we could try to formulate a non-standard logic of ground that does not sanction the logical principles used in the proof of the inconsistency theorem. The analogy between our inconsistency theorem (Theorem 2.9.4) and the Tarski's theorem of the undefinability of truth (Proposition 2.9.5) suggests a neat terminology for these approaches. Analogously to theories of truth [58], we get: *typed theories of partial ground*, which avoid paradox by putting type-restrictions on the relation of partial ground, effectively ruling out self-referential sentences like the one in the proof; *untyped theories of partial ground*, which avoid paradox by restricting the principles of partial ground; and finally *non-classical theories of partial ground*, which avoid paradox (or: triviality) by abandoning classical logic in favor of alternative logics.

The results of §2.6 point in the direction of a typed theory of partial ground. Indeed our theories $PGA_{<\alpha}$, for $1 \leq \alpha \leq \epsilon_0$, *are* typed theories of partial ground: the axiom G_4^α is effectively a typing axiom. Moreover, we have shown that the theories $PGA_{<\alpha}$, for $1 \leq \alpha \leq \epsilon_0$, are consistent (Corollary 2.8.10). Thus, there is good evidence that a typed solution to the new puzzle of ground works. A natural direction to take from here would be to extend Tarski's truth-theoretic hierarchy (cf. §2.6) to a truth- and *ground-theoretic* hierarchy. We would simply type the ground predicate to get a family of predicates: $\lhd_1, \lhd_2, \ldots$. We would end up with a doubly-typed theory: one typing with regard to truth and one typing with regard to ground. The results from §2.6 suggest that this theory will turn out consistent and this provides further support for the claim that we should use typing in the context of theories of ground. But carrying out the details of this proposal is beyond the scope of this chapter.[54]

Before we close this section, we would like to point out one important philosophical consequence of Theorem 2.9.4. Intuitively, applications of the ground predicate to sentences containing the ground predicate are statements about "the grounds of ground": we need such statements to say in virtue of what the truth of a statement of the form $s \lhd t$, for terms s and t, holds. Philosophically, this is an important issue, as we will discuss in the next chapter. But Theorem 2.9.4 shows that we cannot easily address "the grounds of ground" in our predicational framework. Moreover, as we've just said, if we wish to allow the ground predicate to apply to sentences involving the ground predicate, we presumably have to involve heavy technical machinery, as in §2.6. This is a major drawback of the present approach and one of the reasons why we will ultimately abandon it for the purpose of this paper. Statements about the "grounds of ground" naturally come up on our approach to essence. Remember that we wish to say that an essential property of a thing is a property that is grounded in the identity of the thing. But then, if we ask ourselves in virtue of what a property is essential to a thing, the natural answer would be to say that a property is essential to a thing in virtue of being grounded in the identity of the thing. In other words, that a property is grounded in the identity of a thing is grounded in the identity of the thing. But this is a statement about the grounds of

[54]There is also some great intuitive appeal to an untyped theory of partial ground. A suggestion for how this might be achieved can be found in Appendix A.

ground, and as we've just seen, we can't easily formalize such statements in our predicational theory of ground.

2.10 Conclusion

In this chapter we've developed a new and (so we believe) exciting approach to theories of ground. We've shown that simply by formulating the generally accepted principles of partial ground in a predicational framework, we get a natural, both mathematically and philosophically interesting, and most importantly *consistent* theory of partial ground. Moreover, the methods we've used to develop our theory and to address philosophical questions arising from it, especially in §2.6, stem from axiomatic theories of truth. This supports our plea for collaboration between truth-theorists and ground-theorists. We conjecture that a lot of new and hopefully exciting work can be done in this direction. In this chapter, we've only begun to scratch the surface, but we never know what the future holds.

As we've pointed out in §2.3 there are good philosophical reasons for working in a predicational framework. In addition to this we might say that, given the philosophical role that metaphysical ground is supposed to play, the approach is particularly natural. In effect, metaphysical ground is supposed to one of the fundamental building blocks of reality. As Fine [42, p.80] puts it: the relation of ground is supposed to play an important role in "holding up the edifice of metaphysics". But if metaphysical ground is a fundamental concept, then our only recourse for developing a proper theory of ground seems to be axiomatic: we simply postulate as axioms the principles of ground that we hold to be true by reflections on the concept of ground. But, at least for the present purpose, the approach suffers from two major drawbacks: First, at least in its present state, the approach only works for the relation of *partial* ground (see Footnote 2 of this chapter). But when we wish to view the essential properties of a thing as the properties grounded in its identity, we're presumably talking about *full* ground. But when we approach full ground using our ground predicate, we face problems.

Full ground, remember, is the relation of one truth holding wholly in virtue of a possible plurality of others (see Footnote 2 again). Philosophically speaking, it is natural to suppose that the relation of full ground is more fundamental than the relation of partial ground. As Fine [42, p. 50] argues, we can define partial ground in terms of full ground: We simply say that one truth partially grounds another iff the former truth possibly together with some others fully grounds the latter truth. But conversely, Fine argues, it is not possible to define full ground in terms of partial ground. His argument runs as follows: Let φ and ψ be two true sentences. Then, intuitively, the truth of $\varphi \vee \psi$ holds both fully in virtue of the truth of φ and fully in virtue of the truth of ψ. Thus, both the truth of φ and the truth of ψ are full grounds of the truth of $\varphi \vee \psi$. Consequently, the two truths are also partial grounds of the truth of $\varphi \vee \psi$. Now consider the truth of $\varphi \wedge \psi$. Intuitively, the truth of $\varphi \wedge \psi$ holds partially in virtue of the truths of φ and ψ, but not wholly in virtue of either truth. Thus, the truths of φ and ψ are partial grounds, but not full grounds of the truth of $\varphi \wedge \psi$. We have the following scenario: the truths of $\varphi \vee \psi$ and $\varphi \wedge \psi$ have the exact same partial grounds, but different full grounds. Thus, it is unclear how we should define the full grounds of a truth solely based on its partial grounds.

If we take full ground to be more fundamental than partial ground, it is natural to think that we should develop an axiomatic theory of full ground that proves our theory of partial ground as a sub-theory. Full ground is what Correia [24, p. 255] calls a *many-to-one* relation: it is the relation of one truth holding in virtue of a possible *plurality* of others. Now, how should we reflect this fact syntactically? A first approach would be to stick to a binary ground predicate and represent the possible pluralities of truths as *sets* of sentences. But this approach only carries so far. Using the technique of Gödel numbering, we can only represent *finite* sets of sentence, but not *arbitrary* sets of sentences.[55] But in many cases, the plurality of full grounds is intuitively infinite. Think for example of the truth of $\forall x(S(x) \neq \overline{0})$. Intuitively, the truth holds wholly in virtue of all the truths of $S(\overline{0}) \neq \overline{0}, S(\overline{1}) \neq \overline{0}, \ldots$ taken together. Moreover, in the context of our notion of essential properties as properties grounded in identity, we'd want to say that it is, for example, essential to a thing that it is different from everything else. In other words, we'd want to say for all x and for all $y \neq x$, that $y \neq x$ is grounded in the identity of x. But there might be infinitely other things and so the grounds of the essentialist claim for an object a would involve $a \neq b, a \neq c, a \neq d, \ldots$. For this reason, it seems that we need to adopt *multigrade predicates* in the style of [105], which take arbitrary sequents of terms as arguments. If we let $\blacktriangleleft$ be such a multigrade predicate for the relation of strict full ground, we can formalize the first example as:

$$S(\overline{0}) \neq \overline{0}, S(\overline{1}) \neq \overline{0}, \ldots \blacktriangleleft \forall x(S(x) \neq \overline{0}),$$

while we can at the same time write:

$$\varphi \blacktriangleleft \varphi \vee \psi$$

[55] This follows immediately from Cantor's theorem, which entails that the set $\wp(\mathcal{L})$ of all sets of sentences of $\mathcal{L}$ has a strictly bigger cardinality than $\mathcal{L}$ itself. Note that the language $\mathcal{L}$ of Peano arithmetic is countable, i.e. it has the same cardinality of as the natural numbers. Now assume that there is a coding function that injectively maps every set of sentence $\Gamma \subseteq \mathcal{L}$ to a unique code $\#\Gamma \in \mathbb{N}$. This would mean that the cardinality of $\wp(\mathcal{L})$ is less than or equal to the cardinality of $\mathbb{N}$. But this would mean that the cardinality of $\wp(\mathcal{L})$ is less than or equal to the cardinality of $\mathcal{L}$, which is impossible. Thus there is no such coding function.

to say that the truth of φ fully grounds the truth of $\varphi \vee \psi$. Multigrade predicates, however, mean a significant deviation from standard logic to infinitary logic. It would be a non-trivial fact that the results of this chapter still apply in this context and working out the details is (far) beyond the scope of the present chapter.[56]

Second, as we've pointed out in §2.9.1, the present approach cannot easily be adapted to account for the "grounds of ground" and this is a serious obstacle for developing a theory of essence in the present framework. Together, the two problems—failure to account for full ground and failure to account for the "grounds of ground"—lead us to abandon the predicational approach for the purposes of this dissertation. Thus, in the context of this dissertation, the result of the chapter is negative: predicational theories of ground are not developed well enough to provide a natural framework for theories of essence. But in the bigger scheme of things, we hope to have shown that a predicational approach to ground is a new and exciting field of research that will hopefully find a natural home in the logic and metaphysics of ground in the future.

[56]Another infinitary issue that pops up in the context of full ground is that Fine [42] argues that we need an infinitary *totality (multigrade-)predicate* T that applies to a sequence of terms $t_1, t_2, \dots$ iff the denotations of $t_1, t_2, \dots$ make up the domain of discourse. This predicate, so Fine, is needed to properly account for the full grounds of the truths of quantified statements. Once we've moved to a multigrade setting, accommodating such a predicate will be relatively straight-forward, but the step to multigrade predicates is, as we've just pointed out, non-trivial.

Chapter 3

The Full Logic of Worldly Ground

3.1 Preface

In this section, we'll develop our preferred semantic framework for the purpose of this dissertation: to explicate essential properties as properties grounded in identity. The framework traces back to the work of Fraassen [53], who developed an intuitively plausible semantics for the logic of *first-degree entailment* in terms of facts verifying and falsifying sentences [1]. This semantics was recently revived by Fine [42, 44, 38, 45] and brought to bear on a range of philosophically interesting issues, ranging from counterfactuals to intuitionistic logic. What's most important for the present purpose is that the framework can provide a semantics for a worldly conception of ground. This has been shown by Fine [42], who gave semantic clauses for a ground operator that cannot be iterated. In this chapter, we'll extend this semantics so that it gives truth-conditions for iterated applications of the ground operator. This semantic framework, or rather a modified version thereof, will be the one we'll use to explicate the notion of essential properties as properties grounded in the identity of things.

3.2 Introduction

This chapter is about the logic of ground. The aim is to develop a system for the logic of ground, which treats ground as an iterable operator in the sense that it can also be *iterated*. Developing a logic of ground requires us to do three things:

(i) we need to give a *grammar of ground*, which defines the logical form of sentences of ground;

(ii) we need to give a *semantics of ground*, which defines truth and logical consequence for sentences of ground; and

(iii) we need to give a *proof-theory of ground*, which defines the logically valid inferences between sentences of ground.

This already sets the plan for the chapter. But first, we'll give some background and motivation.

3.3 Background and Motivation

3.3.1 The Logic of Ground

Fine [42] defines *ground* as "the intuitive notion of one thing holding in virtue of another" (p. 37).[1] He also gives a paradigmatic example of ground. Assume there's a ball that is both red and round. Then, according to Fine [42, p. 37], the following is true:

(1) The fact that the ball is red and round obtains in virtue of the fact that the ball is red and the fact that the ball is round.

Thus, ground is not a binary relation on the facts, as Fine's definition might suggest, but rather it is a *multiary relation* on the facts: a relation that holds between one fact "on the one side" and many facts "on the other side" [105].[2] In the terminology of Correia [24], ground is *many to one*. If the relation of ground holds between some facts, if one fact holds in virtue of others, we say in ground-theoretic parlance that the latter facts *ground* the former. Thus ground-theoretically speaking, (1) says that the facts that the ball is red and that the ball is round ground the fact that the ball is red and round. We call a sentence like this, which says that the relation of ground holds between some facts, a "sentence of ground."

The *logic of ground* is concerned with which sentences of ground follow from which other sentences of ground—it is concerned with the consequence relation between sentences of ground. In recent years, a great deal progress has been made in the logic of ground. Correia [24, 25] and Fine [42, 44] give very developed systems for the logic of ground, which include both a syntax

[1]For an (opinionated) introduction to ground, see [28]. For an overview of the recent literature on ground, see [21, 133, 118].

[2] Ground in this sense is what Fine [42, p. 50] calls *full ground*: the relation of one fact wholly obtaining in virtue of others. There is also a binary relation of ground that Fine [42, p. 50] calls *partial ground*: the relation of one fact obtaining partially in virtue of another. In this chapter, we'll focus on full ground.

and a semantics of ground.[3] But their systems have an important restriction: they don't allow for *iterated sentences of ground*, where one "in virtue of" occurs within the context of another. They only consider *simple sentences of ground*, where there are no occurrences of "in virtue of" in the context of others. Thus we may call their systems *logics of simple ground*. Litland [88, p. 131–78] and Litland [87] gives a system that accommodates iterated sentences of ground: he gives what we may call a *logic of iterated ground*. But also Litland's system has two important limitations: First, Litland's system is purely syntactic: it is a proof-theory for the logic of (iterated) ground. Moreover, this proof-theory is a higher-order natural deduction system in the sense of Schröder-Heister [126], which, although natural and elegant in the context of the logic of (iterated) ground, is technically quite sophisticated and intuitively hard to grasp. The second limitation is that Litland's system only treats ground in the context of languages without logical operators—it is what Fine [42, p. 54–57] and Fine [44, p. 1] calls a "pure logic of ground." Moreover, it is a non-trivial technical problem to extend Litland's proof-theory to what Fine [42, p. 58–71] calls an *impure logic of ground*, which considers ground in the context of languages with logical operators.

The aim of this chapter is to develop a logic of ground that overcomes the limitations of the previous approaches—the aim is to give what we'll call a *full logic of ground*: an impure logic of both simple and iterated ground. For reasons of simplicity, we shall confine ourselves, for the most part, to the full logic of ground in the context of languages with only the truth-functional connectives, but we shall discuss how the results of the chapter can be extended to the full logic of ground in the context of languages with the quantifiers.

3.3.2 Logics of Iterated Ground

Why should we care about iterated sentences of ground? Here is a philosophical reason: If the relation of ground holds between some facts—if one fact holds in virtue of others—then this is itself a fact: a *grounding fact*. But then we may ask the philosophical question whether there is something that this grounding fact holds in virtue of. And every possible answer to this question takes the form of an iterated sentence of ground. This gives an intuitive meaning to iterated sentences of ground: they are possible answers to questions about what we may call *the grounds of ground*. Under

[3] Schnieder [124] gives a system for the logic of "because," which we can also understand as a logic of ground. But Schnieder only gives a syntax of "because" and not a semantics. Moreover, Schnieder gives a system for a binary use of "because," which corresponds to partial ground. In this chapter, however, we are only interested in the logic of full ground. See footnote 2.

this interpretation, iterated sentences can be used to express the claims of philosophical positions about the grounds of ground.

Let us explore this interpretation of iterated sentences of ground. Metaphysicians have proposed a range of different positions on the grounds of ground: According to a view proposed by Rosen [121, p. 130–33] and Fine [42, p. 74-80] facts about ground are grounded in certain essentialist facts about the concepts involved. Call this the view that *essence grounds ground*. On this view, for example, the fact that a particular conjunctive fact is grounded in its conjuncts is itself grounded in a fact about the nature of conjunction: the fact that it lies in the nature of conjunction that conjunctive facts are grounded in their conjuncts. Think about the red and round ball. On the view that essence grounds ground, if (1) is true, then the following is true:

(2) That the fact that the ball is red and round obtains in virtue of the fact that it is red and the fact that it is round obtains itself in virtue of the fact that it lies in the nature of conjunction that conjunctive facts are grounded in their conjuncts.

On a different view proposed by Bennett [11] and deRosset [32], grounding facts are simply grounded in the grounds that occur in the grounding fact. Call this the view that *grounds ground ground*. Think about the red and round ball again. On the view that grounds ground ground, if (1) is true, then the following is true:

(3) That the fact that the ball is red and round obtains in virtue of the fact that it is red and the fact that it is round obtains itself in virtue of the two facts that the ball is red and that the ball is round.

And on yet a different view proposed by Litland [88, p. 131–78] and Litland [87], it's the grounds together with a certain fundamental fact that ground grounding facts. According to Litland, grounding facts split up into two components: a *factive* component consisting of the grounds and a *non-factive component* consisting of a *non-factive relation of ground*. The notion of ground that we have considered so far is *factive*: it is a relation on the facts. Fine [42, p. 48–50] distinguishes from this a non-factive relation of ground, which he characterizes informally as "obtained from the factive notion by 'rounding out,' in which the possible cases of factive grounding are extended to cases of grounding from impossible antecedents in such a way that the basic principles governing the behavior of ground are preserved" [42, p. 50, quotes in the original]. Litland argues that the non-factive component of a grounding fact is, in a sense, fundamental, since it is not grounded in any other fact. Fine [42, p. 47–48] distinguishes two forms of a fact not being grounded in other facts: First, a fact may be *ungrounded* in the sense that it is not grounded in anything, especially not in any other fact. Second, a fact may be *zero-grounded* in the sense that it is grounded, but in zero facts.

Zero-grounding is thus a form of fundamentality: a fact is zero-grounded if and only if it obtains, but not in virtue of any particular facts. Fine [42] argues that zero-grounding "may be more than an exotic possibility" [42, p. 48]. For example, some necessary truths, such as Socrates being identical to himself, may be best thought of as zero-grounded. For theoretical purposes, it is useful to postulate a *zero-fact* in this context. The zero-fact is a fact that necessarily obtains, but is itself without any content. It allows us to distinguish semantically between ungrounded and zero-grounded facts: a fact is ungrounded if and only if there are no facts that ground it and a fact is zero-grounded if and only if it is grounded by the zero-fact. Putting the factive and the non-factive component of grounding facts together, Litland claims that a grounding fact is grounded in the grounds and the zero-fact. Call this the view that *grounds-plus-zero ground ground*. Thus, on the view that grounds-plus-zero ground ground, if (1) is true, then the following is true:

(4) That the fact that the ball is red and round obtains in virtue of the fact that it is red and the fact that it is round obtains in virtue of the zero-fact together with the two facts that the ball is red and that the ball is round.

These are the main views on the grounds of ground in the literature.[4] Since they are metaphysical views, these views and their consequences for iterated sentences of ground are arguably metaphysically necessary. This has an important consequence: on the view that essence grounds ground, necessarily, if (1) is true, then (2) is true; on the view that grounds ground ground, necessarily, if (1) is true, then (3) is true; and on the view that grounds-plus-zero ground ground, necessarily, if (1) is true, then (4) is true. And, on the view

[4]There is another view on the grounds of ground, which we wont consider in the following. Dasgupta [30] discusses the grounds of ground from the perspective of *ground-theoretic physicalism*: the view that all facts are ultimately grounded in physical facts. Prima facie, grounding facts are a problem for this view because it seems counter-intuitive to say that grounding facts are grounded in physical facts. In light of this problem, Dasgupta defends yet another position on the grounds of ground, according to which *grounding facts are autonomous*: "[grounding facts] are special because they are [...] "not apt for being grounded." It is not that the question of what grounds them is well taken and the answer is "Nothing"; it is, rather, that the question of what grounds them does not legitimately arise in the first place" [30, p. 563]. This suggests that Dasgupta's view is that iterated sentences of ground are meaningless, contrary to what we have claimed above. He argues for his view by attempting to show that it overcomes all the problems of alternative proposals—especially in the light of physicalism. Dasgupta raises many interesting philosophical points, but in this chapter we are not primarily interested in the philosophical merit of the individual positions on the grounds of ground—regardless of the assumption of physicalism. Rather, we are interested in the logic of ground that results from these positions. For this purpose, we take it as a working assumption that iterated sentences of ground are meaningful in the sense sketched above. This working assumption will receive partial vindication from the semantics of ground we'll give later in the chapter: the semantics will give a precise philosophical meaning to iterated sentences of ground.

that essence grounds ground, it's arguably the case that it's possible that (1) is true and (3) is false, that (1) is true and (4) is false, and similarly for the other views. This brings us to the logic of iterated ground. On the standard view of consequence, a sentence φ is a consequence of a set of sentences Γ if and only if it is impossible for all the members of Γ to be true, without φ being true. This means that: on the view that essence grounds ground, (2) is a consequence of (1); on the view that grounds ground ground, (3) is a consequence of (1); and on the view grounds-plus-zero ground ground, (4) is a consequence of (1). And it means that: on the view that essence grounds ground, (3) is not a consequence of (1) and (4) is not a consequence of (1)—and analogously for the other views. The point here is that on different views about the grounds of ground, we get different consequence relations between sentences of ground, and thus different logics of iterated ground.

The aim of this chapter is to explore the logic of iterated ground according to the view that grounds ground ground and to formalize it in a system for the full logic of ground. The results of this paper can be extended with relative ease to the view that grounds-plus-zero ground ground, but not the view that essence grounds ground. The reason is that the view that essence grounds ground requires an account of what Correia [23] calls *generic essence*: a concept of essence or rather essential properties, where we not only have essential properties of objects, but also concepts, properties and so on. To see this, note in our example (2) above, we said that according to the view that essence grounds ground, the grounding fact that the conjunctive fact that the ball is red and round holds in virtue of its two conjunct facts that the ball is red and that the ball is round holds itself in virtue of *the nature of conjunction*. However, as we've said in the introduction, in this dissertation we focus on *objectual essence*: the concept of essence or rather essential properties, where essential properties are properties *of objects* (compare footnote 2, p. 2). And we argue that the view that grounds ground ground gives a natural full logic of ground for this purpose.

3.3.3 Overview

Here is the plan for the rest of the chapter: In §3.4, we'll discuss the syntax of ground and define the (infinitary) languages of ground that we'll be working with. Then, in §3.5, we'll discuss the semantics of these languages. In particular, we'll give semantic clauses that correspond to the view that grounds ground ground. In §3.6, we'll give an (infinitary) proof system for logic of the previous semantics. In the conclusion, we'll sketch how we may obtain a completeness result, but we'll omit the details, since this result would require us to use relatively cumbersome methods from infinitary logic. That's the plan. So, let's get to work.

3.4 The Full Grammar of Ground

3.4.1 The Operator Approach

As we've discussed before in this dissertation, there are two approaches to the semantics of ground in the literature: on the first approach, ground is formalized by means of a relational predicate, while on the second approach ground is formalized by means of an operator [24, p. 253–54, 42, p. 46–48]. Correspondingly, the first approach is called the *predicational approach* to the syntax of ground and the second approach is called the *operator approach* to the syntax of ground. In this chapter, we'll take the operator approach to ground. Thus, we'll take the logical form of a sentence of ground to be:

$$\varphi \textit{ in virtue of } \psi_1, \psi_2, \ldots,$$

where $\varphi, \psi_1, \psi_2, \ldots$ are sentences. The idea is that the sentences flanking the phrase "in virtue of" express the relata of ground. Correspondingly, the intended reading of a sentence of the form φ *in virtue of* $\psi_1, \psi_2, \ldots$ is that what φ expresses holds in virtue of what $\psi_1, \psi_2, \ldots$ express.

3.4.2 Worldly versus Conceptual Ground

As we've discussed in the introduction to this dissertation, there are two competing views of the relata of ground: on the *conceptualist view*, the relata of ground are conceptually individuated *truths*, which we individuate by means of the sentences that express them; and on the *worldly view*, the relata of ground are worldly individuated *facts*, which we individuate by means of the objects, properties, and relations they involve. Here both truths and facts are fact-like entities: entities for which it makes sense to say that they hold or obtain.[5] In this chapter, we will work on the worldly view of the relata of ground. But for reasons of clarity, it is useful to say a bit more about the different logics of ground that result from the two different views.

On the operational approach, we use true sentences to express the relata of ground. But on the two views of the relata of ground, they are different kinds of entities. On the conceptualist view of the relata of ground, we say that a true sentence φ expresses *the truth of* φ. On the worldly view of the relata of ground, in contrast, we say that a true sentence φ expresses *the fact that* φ. The difference between the two views is how we individuate the truths of true sentences and the facts expressed by true sentences.

To illustrate the difference, let's consider disjunctions. On the conceptualist view, we would say that for a true sentence φ,

[5] For more on the distinction between truths and facts, see [41].

- the truth of $\varphi \vee \varphi$ is distinct from the truth of φ.

The reason for making this distinction is that, on the conceptualist view, we would like to say that for all sentences φ and ψ,

- if φ is true, then the truth of $(\varphi \vee \psi)$ holds in virtue of the truth of φ; and
- if ψ is true, then the truth of $(\varphi \vee \psi)$ holds in virtue of the truth of ψ.

In other words, on the conceptualist view, we would like to say that the truth of a true disjunction holds in virtue of the truth of its true disjuncts. If a motivation for this view is wanted, then think of the standard way of determining the truth-value of formulas according to the truth-tables: we determine the truth-value of a disjunction based on the truth values of its disjuncts. Now, by a simple application of the previous principle(s), we get that for all sentences φ,

- if φ is true, then the truth of $(\varphi \vee \varphi)$ holds in virtue of the truth of φ.

But conversely, on the conceptualist view, we don't want to say that for any sentence φ,

- if φ is true, then the truth of φ holds in virtue of the truth of $\varphi \vee \varphi$.

There is both an intuitive and a ground-theoretic justification for this claim on the conceptualist view. For the intuitive justification, think of the truth-tables again: we don't determine the truth-value of a sentence by determining the truth-value of the disjunction of the sentence with itself. For the ground-theoretic justification, first note that most philosophers would accept the following two principle in the logic of ground, regardless of their view of the nature of the relata of ground:

- For no sentence φ and no set of sentences Γ does the truth of φ hold in virtue of the truths of φ and Γ.
- For all sentences φ and ψ and all sets of sentences Γ and Δ, if the truth of φ holds in virtue of the truths of ψ and Γ and the truth of ψ holds in virtue of the truths of Δ, then the truth of φ holds in virtue of the truths of Γ and Δ.

In other words, ground is usually assumed to be irreflexive and transitive.[6] But given these two principles, we can't affirm at the same time that if φ is a true sentence, then the truth of $\varphi \vee \varphi$ holds in virtue of the truth of φ and that the truth of φ holds in virtue of the truth of $\varphi \vee \varphi$. For together

[6] This is, in any case, the standard view of ground. Some authors have challenged this view: Jenkins [66] challenges the claim that ground is irreflexive and Schaffer [122] challenges the claim that ground is transitive. See Litland [86] and Raven [119] for a defense of the standard view against these challenges.

with the two plausible principles, this would lead to a contradiction: first, by the transitivity of ground, we would get that the truth of φ holds in virtue of the truth of φ, and then, by the irreflexivity of ground, we would get a contradiction. Something has got to give. And on the conceptualist view of ground, the plausible candidate is the claim that the truth of φ holds in virtue of the truth of $\varphi \vee \varphi$. But from this observation it follows by the indiscernibility of identicals that the truths of φ and of $\varphi \vee \varphi$ have to be different: since the one truth grounds the other and not vice versa, the two truths have different properties, and thus they are different. The general point here is that, on the conceptualist view, the truths of sentences are very finely individuated, almost as finely as the sentences that express them. Truths of sentences are something like *truth-tropes*.

On the worldly view, in contrast, we would say that for all sentences φ,

- the fact that φ is the same as the fact that $\varphi \vee \varphi$.

In other words, we don't distinguish the fact that is expressed by a sentence and the fact that is expressed by its own disjunction. Intuitively, the idea on the worldly view is that no matter what a sentence φ says about objects, properties, or relations, the sentence $\varphi \vee \varphi$ talks about exactly the same objects, properties, or relations. Indeed, intuitively, the two sentences say exactly the same about whatever they talk about. And thus, on the worldly view, the facts that the two sentences express are the same.

Note, however, that according to the worldly view it's not in general the case that necessarily equivalent sentences express the same facts. To illustrate consider the facts that two times two equals four and that four is an even number. Intuitively,

- the fact that four is an even number holds in virtue of the fact that two times two equals four.

If an argument is wanted, then think of the definition of being even: a natural number n is even iff there is another natural number m such that $2 \times m = n$. And it's plausible to say that an instance of the definiendum holds in virtue of the instance of the definiens—which gives us the claim in question. Now, many philosophers assume that mathematical facts are necessary (compare our Introduction, p. 16). Hence, it's necessary that two times two equals four. But similarly, by Euler's identity, it's necessary that $e^{i\pi} - 1 = 0$. And since, by standard modal logic, any two necessary facts are necessarily equivalent, we have that the fact that two times two equals four is necessarily equivalent to the fact that $e^{i\pi} - 1 = 0$:

- $\Box$(two times two equals four iff $e^{i\pi} - 1 = 0$).

But, intuitively,

- the fact that four is an even number does not hold in virtue of the fact that $e^{i\pi} - 1 = 0$.

Intuitively, the fact that $e^{i\pi} - 1 = 0$ has absolutely nothing to do with the fact that four is an even number—and certainly the one fact doesn't ground the other. Thus, even though on the worldly view the relata of ground are more coarsely individuated than on the conceptualist view, the relata of ground are still relatively fine-grainedly individuated on the worldly view.

Note that so far we haven't given an *identity criterion* for the facts expressed by sentences on the worldly view. But, following Quine's dictum "no entity without identity" [106, p. 23], it would be desirable to have a necessary and sufficient condition for the facts expressed by two sentences to be the same according to the worldly view. In other words, we would like to fill in the dots in:

- For all true sentences φ and ϕ, the fact expressed by φ is identical to the fact expressed by ψ iff ....

All we know so far, is that two true sentences φ and ψ being necessarily equivalent is not sufficient for the facts that they express to be the same. So how should we will the gap in the identity criterion. For this purpose, Correia [24, p. 256–59] introduces the notion of sentences being *factually equivalent*. For sentences φ and ψ, let us write $\varphi \leftrightharpoons \psi$ to say that φ and ψ are factually equivalent. Then, the idea is that:

- For all true sentences φ and ϕ, the fact expressed by φ is identical to the fact expressed by ψ iff $\varphi \leftrightharpoons \psi$.

But this idea only has merit, of course, if we have an understanding of the notion of factual equivalence: all that we have gained so far is a precise terminology for the question at hand. To get a better understanding of the notion of factual equivalence, Correia [24, 26] suggests that the *logic* of factual equivalence is Angell's logic of *analytic equivalence* [2]. Correia [27] develops a semantics for this logic, which we can thus take as the semantics of factual equivalence as well. Correia's semantics, however, is *algebraic*, in the sense that it's based on a class of algebraic structures, which are given by a set of defining equations. This semantics gives us an understanding of the *structure* of the facts and factual equivalence according to the worldly view. But the semantics does not give us an *intuitive* understanding of facts and factual equivalence. The problem here really is that the structures used in the algebraic semantics are given by a set of defining equations, and thus every collection of objects that satisfy these equations can play the role of the facts in this semantics. Put differently, all the properties and relations of the "facts" in a structure of this semantics are only defined in relation to one another, and thus we don't get an independent, intuitive understanding of facts and factual equivalence from this semantics. It is a plausible condition on a semantics, however, that if it is to give us a philosophical understanding of the concepts it applies to, then it should be phrased in terms of independently motivated and ideally previously understood concepts. And it seems that Correia's semantics does not satisfy this condition.

In recent work, Fine [37] proposes a semantic for Angell's logic of analytic equivalence that satisfies our condition. Fine's semantics is phrased in the framework of *exact truthmaker semantics*, which ultimately traces back to Fraassen [53]. The idea of this semantics is that we interpret sentences by means of finely individuated *states (of affairs)*, which can *exactly verify* and *exactly falsify* sentences. For example, on this semantics, there will be the state of the ball being red, the state of the ball being round, and the state of the ball being red and round. And on this semantics, we say that the state of the ball being red exactly verifies the sentence "the ball is red," the state of the ball being round exactly verifies the sentence "the ball is round," and the state of the ball being red and round exactly verifies the sentence "the ball is red and the ball is round." What sets this semantics off from other truthmaker approaches is that the relation of verification is supposed to be *exact*: for example, the state of the ball being red and round is *not* an exact verifier of the sentence "being red," even though it contains (in a mereological sense) an exact verifier of the sentence. Intuitively, the exact verifiers of a sentence are all the states such that if they obtain, they are *directly responsible* for the truth of the sentence. So far, we've only considered actually obtaining states, but on the semantics, there are also possibly obtaining states of affairs. For example, there will also be the state of the ball being blue, the state of the ball being green, and so on. According to the semantics, these states are considered to be *exact falsifiers* of the sentence "the ball is red." Again, intuitively, the exact falsifiers of a sentence are all the states such that if they obtain, they are *directly responsible* for the falsehood of the sentence. Analogously to the case of exact verification, also exact falsification is supposed to be exact: even though the state of the ball being green and round contains (in a mereological sense) an exact falsifier of the sentence "the ball is red," it is not itself an exact falsifier of the sentence. Finally, once we know the exact verifiers and falsifiers of sentences, we can give the truth-conditions for a sentence φ by saying that:

- φ is true iff some exact verifier of φ obtains, and
- φ is false iff some exact falsifier of φ obtains.

We can then put conditions on the assignment of exact verifiers and falsifiers to sentences, such that the resulting logic will be classical: for example, we could demand that for no sentence φ there is an exact verifier of the sentence and an exact falsifier of the sentence, such that it is possible for both states to obtain; or we could demand that for every sentence φ, necessarily, either an exact verifier or an exact falsifier of the sentence obtains. These two conditions together would ensure classicality of the semantics.

In this semantic framework, Fine's analysis of factual equivalence is that two (true) sentences φ and ψ are factually equivalent iff they have the same exact verifiers and falsifiers:

- For all (true) sentences φ and ψ, $\varphi \leftrightharpoons \psi$ iff for all states s, (s is an exact verifier of φ iff s is an exact verifier of ψ) & (s is an exact falsifier of φ iff s is an exact falsifier of ψ).

Fine [37] proceeds to show that this gives us a sound and complete semantics for Angell's logic of analytic equivalence. Thus, on this semantics, we get an identity criterion for facts according to the worldly view:

- For all true sentences φ and ϕ, the fact expressed by φ is identical to the fact expressed by ψ iff for all states s, (s is an exact verifier of φ iff s is an exact verifier of ψ) & (s is an exact falsifier of φ iff s is an exact falsifier of ψ).

Using this criterion, we can say more specifically what the relata of ground according to the worldly view are: according to Fine's semantics, they are effectively pairs of sets of states, which intuitively consist of a set of exact verifiers and a set of exact falsifiers. Together with the intuitive interpretation of this semantics that we've just discussed, this framework gives us an understanding of facts and factual equivalence on the worldly view, which moreover satisfies the intuitive constraint that it should be phrased in terms of independently motivated and ideally previously understood concepts.

In this chapter, we will employ Fine's framework for our semantics of ground. Fine [42, p. 71–74] has shown how we can give truth-conditions for sentences of ground in this semantic framework, and in this paper, we will extend Fine's semantics to cases of iterated ground on the worldly view. The upshot of the discussion for now is that in our language of ground, we will have an operator $\leftrightharpoons$ for factual equivalence, in terms of which we can explain the identity of the relata of ground and which gets the intended interpretation given by Fine's semantics for analytic equivalence.

3.4.3 Weak Ground versus Strict Ground

Fine [42, p. 48–54] distinguishes various of readings of the phrase "in virtue of," or, in other words, various *concepts of ground.* Of the different concepts that Fine distinguishes, two are especially important for the present purpose: the notions of *strict* and *weak* ground.

The crucial issue on which the distinction between strict and weak ground rests, is the question whether ground is irreflexive. As we've said above, ground is usually regarded as an irreflexive and transitive relation. On the worldly view, this means that the following two principles are usually accepted:

Irreflexivity. For no sentence φ and no set of sentences Γ does the fact expressed by φ hold in virtue of the facts expressed by φ and the members of Γ.

Transitivity. For all (true) sentences φ and ψ and all sets of sentences Γ and Δ, if the fact expressed by φ holds in virtue of the facts expressed by ψ and the members of Γ and the fact expressed by ψ holds in virtue of the facts expressed by the members of Δ, then the fact expressed by φ holds in virtue of the facts expressed by the members of Γ and Δ.

The reading of "in virtue of" according to which these two principles are true is is what Fine calls *strict ground*: the notion is strict because it is irreflexive.

In contrast, Fine argues that there is also a coherent notion of ground that is not only not irreflexive, but indeed reflexive in the sense that the notion satisfies the following principle:

Reflexivity. For every (true) sentence φ, the fact expressed φ holds in virtue of the fact expressed by φ.

The reading of "in virtue of" according to which **Reflexivity** and **Transitivity** are true is what Fine calls *weak ground*: the notion is weak because it is reflexive.

According to Fine [42, p. 51–52] the weak notion of ground is often not expressed by means of "in virtue of," but rather by means of phrases of the form

for the fact that φ to hold is for the facts that $\psi_1, \psi_2, \ldots$ to hold,

where $\varphi, \psi_1, \psi_2, \ldots$ are sentences. The idea is that a sentence of this form says that the facts expressed by $\psi_1, \psi_2, \ldots$ are weak grounds for the fact expressed by φ. On this way of expressing ground, it is intuitively plausible that we get the following two principles:

- For every (true) sentence φ, for the fact expressed φ to hold is for the fact expressed by φ to hold.
- For all sentences φ and ψ and all sets of sentences Γ and Δ, if for the fact expressed by φ to hold is for the facts expressed by ψ and the members of Γ to hold and for the fact expressed by ψ to hold is for the facts expressed by the members of Δ to hold, then for the fact

expressed by φ to hold is for the facts expressed by the members of Γ and Δ to hold.

Both principles are intuitively motivated from the intuitive reading of the phrase "for ... to hold is for ___ to hold." Thus, if we take the phrase "for ... to hold is for ___ to hold" to express a concept of ground, then it is the weak concept of ground.

Now Fine claims that we can indeed read the phrase "for ... to hold is for ___ to hold" as expressing a concept of ground. To motivate this, let's consider Hesperus, Phosphorous, and Venus. Since Hesperus, Phosphorous, and Venus are pairwise identical, the following claims are intuitively plausible:

- For the fact that Hesperus is Venus to hold is for the fact that Phosphorous is Venus to hold.
- For the fact that Phosphorous is Venus to hold is for the fact that Hesperus is Venus to hold.

Now, if we ask ourselves what this means in terms of "in virtue of," we might be tempted to accept the following two claims:

- The fact that Hesperus is Venus holds in virtue of the facts that Hesperus is Phosphorous and that Phosphorous is Venus.
- The fact that Phosphorous is Venus holds in virtue of the facts that Hesperus is Phosphorous and that Phosphorous is Venus.

But given **Transitivity**, these two principles spell trouble, for they entail that:

- The fact that Hesperus is Venus holds in virtue of the facts that Hesperus is Phosphorous and that Phosphorous is Venus.

And this is in direct contradiction to **Irreflexivity**. Now, Fine's proposal is to say that in such a case, we should not read "... in virtue of ___" as expressing strict ground, but rather as meaning the same as "for ... to hold is for ___ to hold"—as expressing a concept of weak ground. For note that the following claim is not only innocuous, but indeed intuitively plausible:

- For the fact that Hesperus is Venus to hold is for the facts that Hesperus is Phosphorous and that Phosphorous is Venus to hold.

Thus, if we follow Fine, and read "... in virtue of ___" as meaning the same as "for ... to hold is for ___ to hold," then we don't face a problem. In other words, there is a coherent reading of "in virtue of" as expressing weak ground.

Now, we're in the somewhat uncomfortable situation of having different expressions for ground: one and the same expression ("in virtue of") can express a different concept of ground in different contexts, and we can have different expressions ("in virtue of" and "for ... to hold is for ___ to hold") which can express the same concept. It is time to regiment the syntax of ground to avoid confusion. In the following, we'll use Fine's operators $<$ and $\leq$ to express strict and weak ground respectively. Thus, a sentence of the form

$$\varphi < \psi_1, \psi_2, \ldots,$$

where $\varphi, \psi_1, \psi_2, \ldots$ are sentences, says that the fact expressed by φ holds in virtue of the facts expressed by $\psi_1, \psi_2, \ldots$ in the *strict* sense of ground. And a sentence of the form

$$\varphi \leq \psi_1, \psi_2, \ldots,$$

where $\varphi, \psi_1, \psi_2, \ldots$ are sentences, says that the fact expressed by φ holds in virtue of the facts expressed by $\psi_1, \psi_2, \ldots$ in the *weak* sense of ground. Correspondingly, we'll call $<$ the *strict ground operator* and $\leq$ the *weak ground operator*.[7]

The notion of weak ground is contested in the literature. In particular, Derosset [33] argues that the concept of weak ground is intuitively obscure and only the concept of strict ground is coherent.[8] But for developing a logic and semantics of (iterated) ground, it is useful to have the notion of weak ground at hand, so we should say something in defense of the notion.

Fine [42] points out that the notion of weak ground that he has in mind can be *defined* in terms of the weak notion of ground. In particular, he says that for all sentences $\varphi, \psi_1, \psi_2, \ldots$ we can say that:

- $\psi_1, \psi_2, \ldots \leq \varphi$ iff for all sentences θ and all sets of sentences Γ, if $\psi_1, \psi_2, \ldots, \Gamma < \theta$, then $\varphi, \Gamma < \theta$.

In other words, $\psi_1, \psi_2, \ldots$ weakly ground φ iff $\psi_1, \psi_2, \ldots$ subsume the ground-theoretic role of φ. Under this definition, it is easy to show that weak ground is reflexive and, assuming that strict ground is transitive, so is weak ground. Moreover, the definition single-handedly gives us a better intuitive grasp of the concept of weak ground and discards DeRossett's worries: if we say that strict ground is coherent, then certainly also every notion that can be defined in terms of strict ground is coherent; and since weak ground can be defined in terms of strict ground, we seem to be committed to saying that also the concept weak ground is coherent.

For the purpose of developing a semantics and logic of ground it is useful to take the concept of weak ground as primitive. However, as Fine [42, p. 52]

[7] Technically speaking, we'll take the arguments to the left of the operators $<$ and $\leq$ to be *sets* of sentences. By standard convention, we'll omit outermost set-brackets. So instead of $\{\psi_1, \psi_2, \ldots\} < \varphi$, we write $\psi_1, \psi_2, \ldots < \varphi$, and analogously for $\leq$. And if Γ is a set of sentences, we write $\Gamma, \varphi < \psi$ instead of $\Gamma \cup \{\varphi\} < \psi$, and analogously for $\leq$.

[8] But see [66] for an argument that ground is reflexive in at least some cases.

points out, it is also possible to define strict ground in terms of weak ground, by saying for all sentences $\varphi, \psi_1, \psi_2, \ldots$ that:

- $\psi_1, \psi_2, \ldots < \varphi$ iff $\psi_1, \psi_2, \ldots \leq \varphi$ & for no ψ_i there is a set of sentences Γ such that $\varphi, \Gamma \leq \psi_i$.

In other words, strict ground can be defined as weak ground that cannot be reversed. We can easily show that according to this definition, strict ground is indeed irreflexive and, given the assumption that weak ground is transitive, we can show that strict ground is also transitive. Thus, even if we take weak ground as primitive, rather then strict ground, we don't lose the concept of strict ground: it can be recovered via the above definition.

Taking the concept of weak ground as primitive has the benefit that we don't need an extra symbol $\leftrightharpoons$ for factual equivalence in our language: we can *define* this relation in terms of weak ground. The idea is that we can take factual equivalence to be weak ground in both directions. More formally, we can say:

- For all sentences φ and ψ, $\varphi \leftrightharpoons \psi$ iff $\varphi \leq \psi$ & $\psi \leq \varphi$.

In due course, we will show semantically that this definition captures indeed the concept of factual equivalence that we defined above. The upshot of our discussion here is if we take weak ground as our primitive concept of ground, and correspondingly the weak ground operator $\leq$ as our primary means for expressing ground, then we can recover both strict ground and factual equivalence—given that we can recover the above definitions in our framework.

3.4.4 Languages of Ground

The last thing to do in this section is to properly define the languages that we'll be working with. For reasons of simplicity, we'll work with relatively weak languages that have individual constants, predicates, and the truth-functional connectives, but no function symbols, variables, or quantifiers. It is relatively easy, though tedious, to extend our semantics to accommodate function symbols; we mainly omit them for reasons of technical convenience. Variables and quantifiers, however, present more serious technical difficulties. The problem is that in order to accommodate the quantifiers we effectively need to use to a relatively strong infinitary methods—something that we'd like to avoid as much as possible for reasons of perspicuity.[9] Using such

[9]To be more precise, one of the problems is that the logic of a language with finite quantification and countably infinite conjunctions and disjunctions, usually indicated by writing $\mathcal{L}_{\omega_1,\omega}$, is the only logic which is, as Bell [10] puts it, "reasonably well behaved" (§5). But at this level we already lose standard meta-logical results, such as the compactness theorem. In this logic, we only get a "surrogate" for compactness, known as the *Barwise*

strong infinitary logic is not in itself a problem, and as we'll see, there are good reasons for using infinitary logics in the context of ground. However, to illustrate the semantics we have in mind here, "going strongly infinitary" is more trouble than its worth, if we may put it in such a lax way.

We specify a language of ground by specifying its *vocabulary*:

Definition 3.4.1. *Let $\mathcal{L}$ be a language of ground. The* logical vocabulary *of $\mathcal{L}$ consists of:*

(i) the identity symbol $=$*;*

(ii) the truth-functional connectives*:* $\neg, \bigvee$, *and* $\bigwedge$;

(iii) the weak ground operator*:* $\leq$*;*

(iv) the auxiliary symbols*:* (,), *and* ,.

The non-logical vocabulary *of $\mathcal{L}$ consists of:*

(iv) a set $\mathcal{C}$ of individual constants; and

(v) a set $\mathcal{P}$ of predicate symbols together with a function $ar : \mathcal{P} \to \mathbb{N}$ that assigns to every predicate symbol $P \in \mathcal{P}$ a natural number $ar(P)$.

For a predicate symbol $P \in \mathcal{P}$, we call $ar(P)$ its arity*. Consequently, if $ar(P) = n$, for $P \in \mathcal{P}$ and $n \in \mathbb{N}$, we say that P is an n-*ary *predicate symbol.*

We can then define the *sentences* of a language of ground in the usual recursive way:

Definition 3.4.2. *Let $\mathcal{L}$ be a language of ground, then the class of* sentences *of $\mathcal{L}$ is defined by saying that:*

(i) if $c_1, c_2 \in \mathcal{C}$, then $c_1 = c_2 \in X$;

(ii) if $P \in \mathcal{P}$ with $ar(P) = n$ and $c_1, \ldots, c_n \in \mathcal{C}$, then $P(c_1, \ldots, c_n) \in \mathcal{L}$;

(iii) if $\varphi \in \mathcal{L}$, then $\neg\varphi \in \mathcal{L}$;

(iv) if $\Gamma \subseteq \mathcal{L}$, then $\bigwedge \Gamma, \bigvee \Gamma \in \mathcal{L}$;

(v) if $\Gamma \subseteq \mathcal{L}$ and $\varphi \in \mathcal{L}$, then $(\Gamma \leq \varphi) \in \mathcal{L}$, and

---

Compactness Theorem. Roughly, this theorem states that compactness holds in the logic for all *admissible sets* of formulas. The definition of this concept of admissibility, however, is relatively complicated. Moreover, in even stronger infinitary logics, giving reasonably well behaved deductive systems becomes increasingly difficult. And as soon as we allow infinitary quantification, completeness has to be abandoned as well. Because of these complications, we will not go into more complex infinitary logics than we'll have to in this chapter. Our logic will, in any case, be formulated in languages with arbitrarily infinitary conjunctions and disjunctions, but since we're not interested into strong meta-results, this does not create many complications. For an overview of the issues that arise in infinitary logics, see [69, 68, 56, 10].

(vi) nothing else is a member of $\mathcal{L}$.[10]

In the following, we'll denote the set of sentences of $\mathcal{L}$ simply by $\mathcal{L}$: for all expressions φ, $\varphi \in \mathcal{L}$ means φ is a sentence of $\mathcal{L}$.

Note that in Definition 3.4.2, we use $\bigwedge$ ('conjunction') and $\bigvee$ ('disjunction') as operators on *sets* of formulas, rather than formulas themselves. Since these sets may be infinite, clause (iv) of Definition 3.4.2 allows us to form *infinite conjunctions and disjunctions*. We introduce the following notational conventions: For finite sets of formulas, instead of $\bigvee\{\varphi_1, \ldots, \varphi_n\}$, we'll also write $\varphi_1 \vee \ldots \vee \varphi_n$, and analogously for $\bigwedge$. And for indexed (possibly infinite) sets of formulas, we'll also write $\bigwedge_{i \in I} \varphi_i$ instead of $\bigwedge\{\varphi_i \mid i \in I\}$, and analogously for $\bigvee$. Note that since for *all* sets $\Gamma \subseteq \mathcal{L}$ there is a conjunction $\bigwedge \Gamma$ and a disjunction $\bigvee \Gamma$, in particular, there is also an *empty conjunction* $\bigwedge \emptyset$ and an *empty disjunction* $\bigvee \emptyset$. Traditionally, these two can play the role of the *verum* $\top$, in the case of $\bigwedge \emptyset$, and the *falsum* $\bot$. To motivate this note that a conjunction of the form $\bigwedge \Gamma$ is intuitively true iff all the members of Γ are true. But this means that $\bigwedge \emptyset$ is *trivially true*: it has no members and thus all of its members are always true. In contrast, a disjunction of the form $\bigvee \Gamma$ is intuitively true iff some members of Γ are true. But this means that $\bigvee \emptyset$ is *trivially false*: it has no members, so it's impossible for some of its members to be true.

Treating infinite conjunctions and disjunctions in this way is familiar from infinitary (propositional) logic. The reason why we do so in the context of languages of ground is that they allow us to recover the definitions of factual equivalence and strict ground in terms of weak ground in our language:

Definition 3.4.3. *Let $\mathcal{L}$ be a language of ground. Then for all formulas φ, ψ and all sets of formulas Γ:*

(i) $\varphi \fallingdotseq \psi =_{def} (\varphi \leq \psi) \wedge (\psi \leq \varphi)$*; and*

(ii) $\Gamma < \varphi =_{def} (\Gamma \leq \varphi) \wedge \bigwedge\{\neg(\varphi, \Delta \leq \psi) \mid \Delta \subseteq \mathcal{L}, \psi \in \Gamma\}$.

Thus, $\varphi \fallingdotseq \psi$ and $\Gamma < \varphi$ are mere syntactic *abbreviations* for the corresponding definientia in Definition 3.4.3. These syntactic definitions simply reflect, in a syntactic way, the informal definitions of factual equivalence and strict ground in terms of weak ground that we've discussed in §3.4.3. We'll show in due course that they are indeed semantically adequate, but for now note one perk of treating factual equivalence and strict ground in this way: we only need to discuss the semantics and logic of weak partial ground (and the logical vocabulary, of course) and, assuming that their definitions are correct, the logic of factual equivalence and strict ground will be already be taken care of this and the infinitary background logic. This gives us a good

[10]Note that the collection of formulas of $\mathcal{L}$ *class sized*. But this doesn't create any serious problems for what we do in the following.

motivation for using infinitary logic in the context of the logic of ground.[11]

Finally, if $\mathcal{L}$ is a language of ground, then we define the language $\mathcal{L}_{\setminus\leq}$ to the sub-language of $\mathcal{L}$ that is defined just like $\mathcal{L}$, except that we omit clause (v). Thus, effectively $\mathcal{L}_{\setminus\leq}$ is just $\mathcal{L}$ without sentences that have the ground operator in them. It is easily checked that $\mathcal{L}_{\setminus\leq} \subseteq \mathcal{L}$. We define this restricted language merely for technical convenience in the exposition of various results.

3.5 The Full Semantics of Ground

In §3.4.2, we have already sketched the idea of the semantic framework that we'll use for the semantics of ground. The idea is that instead of the notion of a sentence being true at a possible world, which underlies the Kripke's possible-worlds semantics, we now have the notion of a possible state exactly verifying a sentence as the fundamental notion of our semantics. In short, possible worlds are replaced with possible states. As Fine eloquently puts it:

> *[T]he pluriverse of possible worlds is replaced with a space of possible states—the monolithic blobs shatter into myriad fragments.* [38, p. 233]

Just as possible worlds intuitively correspond to the ways the world could have been like, possible states intuitively correspond to the ways things could have been like. And just like there intuitively is a possible world for every way the world could have been like, there are possible states for every way some objects could have been like. Thus, since the ball could have been green, blue, red, and so on, there will be possible states of the ball being green, blue, red, and so on. And just like a sentence about the world is true at a possible world iff what world is like what the sentence says about the wold, a sentence about some objects is exactly verified by a state iff the state corresponds exactly to what the sentence says about the objects. Thus, for example, the sentence "the ball is red" is intuitively exactly verified by the possible state of the ball being red, but not by the states of the ball being green or the ball being blue. And just like a sentence is actually true on the possible worlds approach iff what the sentence says about the world is the case at the actual world, on the exact verification approach a sentence is true iff some state that exactly verifies the sentence actually obtains. Thus, assuming that the ball is actually red, the sentence "the ball is red" is true.

[11] The fact that we need some infinitary devices to properly formulate a logic of ground is implicitly acknowledged by Fine [42]: when he formulates his logic of *conceptualist* ground, he uses infinitary devices that amount to infinite conjunction and disjunction. However, Fine's infinite devices are built into the structure of the logic, as it were, and here we bring the infinitary issues to light more explicitly by simply using infinite conjunctions and disjunctions.

However, this analogy only carries so far: the state based semantics diverges from the possible in some quite substantial ways. First, the notion of exact verification is quite sensitive to the concept of source. For example, the possible state of the ball being red and round does *not* exactly verify the sentence "the ball is red," since it does not *exactly* correspond to what the sentence says about the ball. Intuitively, the exact verifiers of a sentence are the possible states that are *directly responsible* for the sentence being true, whenever they obtain. And only the possible state of the ball being red seems to fit the bill. In contrast, on the possible worlds approach, the sentence "the ball is red" is true at any possible world where the ball is both red and round and it is true at any world where the ball is just red. In other words, the notion of a sentence being true at a world is less discriminating than the notion of a state exactly verifying a sentence.

Second, on the state based semantics we don't only care what makes a sentence true, but also what makes it false: for this reason, next to the notion of exact verification, we also have the notion of exact *falsification.* Just like the exact verifiers of a sentence are the possible states that are directly responsible for the sentence being true, whenever they obtain, the exact falsifiers of a sentence are the possible states that are directly responsible for the sentence being false, whenever they obtain. So, for example, the possible states of the ball being green, the ball being blue, and so on are all exact falsifiers of the sentence "the ball is red." And just like the notion of exact verification, the notion of exact falsification is sensitive to source. The state of the ball being green and round, for example, is *not* an exact falsifier of the sentence "the ball is red"—only the simple states the ball being green, the ball being blue, and so on are. On the possible worlds approach, in contrast, a sentence about the world is false at a possible world iff what the sentence says about the world is not the case at the possible world. So, for example, the sentence "the ball is red" is false at any world where the ball is green, but also at every world where the ball is green and round.

And third, possible states are relatively *local* and they can be *incomplete*: a possible state is a state of some objects being a certain way—and only of those objects being that way. States may nevertheless be (mereologically) *fused*: since there is a state of the ball being red and the state the cup being empty, there is a possible state of the ball being red and the cup being empty. This state is simply the *fusion* of those states. If we keep on fusing states, we will eventually end up with a complete state, which roughly corresponds to a possible world, but there are also incomplete states. Possible worlds, in contrast, are intuitively *complete*: for everything that can either be the case or not, at every world it is either the one way or another. Moreover, it is plausible to say that two possible worlds are identical iff everything that is the case at the one is also the case at the other and vice versa. In other words, possible worlds are individuated by what's the case at them. Thus, at any two distinct possible world, something will be the case at the one that is not the case at the other. For this reason, it does not make sense to fuse possible worlds: since everything is either one way or the other at every possible world, if we combine two possible worlds, we'll end up with an inconsistency. And at least as long as we want to assume that possible worlds are consistent, this is impossible. The point here is that states are more like parts of what is the case at possible worlds, they are the local facts that obtain at worlds, while worlds are determined by all that is the case at the world. This is the sense in which Fine says that "the monolithic blobs shatter into myriad fragments" [38, p. 233].

3.5.1 State Spaces and Interpretations

It is time to make the intuitive idea of exact truthmaker semantics formally precise. For this purpose, Fine [42, 44, 38] introduces what he calls *state spaces*. Here we will slightly modify Fine's state spaces, to accommodate talk of objects in this framework. For this purpose, we introduce *domains* of objects into the framework. And since we wish to allow for some objects to not actually exist, we'll have two domains actually: an *inner domain* of actual objects and an *outer domain* of possible objects.[12] We get:

Definition 3.5.1. *A* state space $\mathcal{S}$ *is an ordered 5-tuple* $(D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ *which consists of:*

(i) a non-empty set $D_\Diamond$ *of* possible objects*;*

(ii) a non-empty set $D_@ \subseteq D_\Diamond$ *of* actual objects*;*

(iii) a non-empty set $S_\Diamond$ *of* possible states*;*

(iv) a non-empty set $S_@ \subseteq S_\Diamond$ *of* actual states*;*

(v) $\prod : \wp(S_\Diamond) \to S_\Diamond$ *is an operation of* state fusion*, such that:*

(a) for all $X \subseteq S_@, \prod X \in S_@$*;*

(b) for all $X \subseteq S_\Diamond$*, if* $\prod X \in S_@$*, then* $X \subseteq S_@$*;*

(c) for all states $s \in S_\Diamond, \prod\{s\} = s$ *('idempotence'); and*

(d) for all indexed families $(X_i \subseteq S_\Diamond)_{i \in i}$ *of states,*

$$\prod\{\prod X_i \mid i \in I\} = \prod \bigcup \{X_i \mid i \in I\} \text{ ('commutativity').}$$

This definition gives us a precise meaning to the informal framework that we've laid out above.

[12] This ties in with our assumptions from the introduction that we wish to subscribe to **Contingentism** and that the background logic for our form of essentialism is a (negative) free logic.

In the context of a state space, we can define a few useful notions for exact verifier semantics:

Definition 3.5.2. *Let $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ be a state space, then we define:*

(i) *the* world state @ *by saying that* $@ =_{def} \prod S_@$; *and*

(ii) *the* zero-state λ *by saying that* $\lambda =_{def} \prod \emptyset$.

It follows immediately from clause (v.a) of Definition 3.5.1 that for a state space $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ that both $@ \in S_@$ and $\lambda \in S_@$. The zero-state λ can play the role of the zero fact that we mentioned in §3.5.1. The world state @, on the other hand, corresponds to the fusion of all the states that are intuitively the case at the actual world.

Definition 3.5.3. *Let $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ be a state space, then we say that a set $X \subseteq S_\Diamond$ is* closed *iff for all non-empty $Y \subseteq X$, we have $\prod Y \in X$.*

It follows immediately from clause (v.a) of Definition 3.5.1 that $S_@$ is closed.

Next we define the notion of an *interpretation* for a language of ground:

Definition 3.5.4. *Let $\mathcal{L}$ be a language of ground and $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space. An* interpretation *for $\mathcal{L}$ in $\mathcal{S}$ is an ordered triple $\mathcal{I} = (\delta, v^+, v^-)$, which consists of:*

(i) *a* denotation function *δ that assigns to every individual constant $c \in \mathcal{C}$ a* denotation *$\delta(c) \in D_\Diamond$;*

(ii) *a* verifier assignment *v^+ which assigns to every n-ary predicate symbol $P \in \mathcal{P}$ a n-ary function $v^+(P) : D_\Diamond^n \to \wp(\mathcal{S}_\Diamond)$ which in turn assigns to every n-tuple $(d_1, \ldots, d_n) \in D_\Diamond^n$ a closed set of states $v^+(P)(d_1, \ldots, d_n) \subseteq \mathcal{S}_\Diamond$;*

(iii) *a* falsifier assignment *v^- which assigns to every n-ary predicate symbol $P \in \mathcal{P}$ a n-ary function $v^+(P) : D_\Diamond^n \to \wp(\mathcal{S}_\Diamond)$ which in turn assigns to every n-tuple $(d_1, \ldots, d_n) \in D_\Diamond^n$ a closed set of states $v^+(P)(d_1, \ldots, d_n) \subseteq \mathcal{S}_\Diamond$.*

We say that an interpretation $\mathcal{I} = (\delta, v^+, v^-)$ for a language of ground $\mathcal{L}$ in a state space $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ is negatively adequate *iff*

(a) *for no n-ary predicate symbol $P \in \mathcal{P}$ there are $d_1, \ldots, d_n \in D_@$ such that:*

$$\text{both } v^+(P)(d_1, \ldots, d_n) \cap S_@ \neq \emptyset \text{ and } v^-(P)(d_1, \ldots, d_n) \cap S_@ \neq \emptyset$$

(b) *for all n-ary predicate symbols $P \in \mathcal{P}$ and for all $d_1, \ldots, d_n \in D_@$ we have that:*

$$\text{either } v^+(P)(d_1, \ldots, d_n) \cap S_@ \neq \emptyset \text{ or } v^-(P)(d_1, \ldots, d_n) \cap S_@ \neq \emptyset$$

(c) for all n-ary predicate symbols $P \in \mathcal{P}$ and for all $d_1, \ldots, d_n \in D_\Diamond$ we have that $v^+(P)(d_1, \ldots, d_n) \cap S_@$, only if $d_1, \ldots, d_n \in D_@$.

In the following, we'll assume that all interpretations are negatively adequate.

It might be helpful to give an intuitive interpretation to the notions we have just defined. So, let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. Now, the interpretation of the denotation function δ should be clear: it assigns denotations to individual constants. Note that we have not postulated that the denotation $\delta(c)$ of a term $c \in \mathcal{C}$ has to be an actual object—constants may denote non-existing objects.

For the intuitive interpretation of v^+ and v^- let's first note that intuitively n-ary predicate symbols express n-ary relations. What the verifier assignment v^+ does is that it assigns to an n-ary predicate symbol $P \in \mathcal{P}$ another function which tells us for every n-tuple $(d_1, \ldots, d_n) \in D_\Diamond^n$ of possible objects the exact conditions for these object to stand in the relation expressed by P. The idea is that objects $d_1, \ldots, d_n \in D_\Diamond^n$ stand in the relation expressed by P iff some member of $v^+(P)(d_1, \ldots, d_n)$ is an actual state—iff the state is a member of $S_@$. Thus, we may think of $v^+(P)(d_1, \ldots, d_n)$ as a disjunctive list of exact *criteria* for the objects $d_1, \ldots, d_n$ to exemplify the relation expressed by P. The condition that $v^+(P)(d_1, \ldots, d_n)$ be closed amounts to the intuitive claim that these criteria may be overdetermined: if we have a set $X \subseteq v^+(P)(d_1, \ldots, d_n)$ of criteria for exemplification, then also their fusion $\prod X$ is a criterion of exemplification. So, for example, if we take the predicate symbol "x is colored" then we would want to assign this predicate the function which maps every object $d \in D_\Diamond$ to all the possible set of states of the object is colored: the state of d being blue, the state of d being red, and so on.

Similarly, v^- gives us what we may call the *anti-exemplification criteria* for properties: what v^- does is that it assigns to an n-ary predicate symbol $P \in \mathcal{P}$ another function which tells us for every n-tuple $(d_1, \ldots, d_n) \in D_\Diamond^n$ of possible objects the exact conditions for these object *not* to stand in the relation expressed by P. The idea is again that objects $d_1, \ldots, d_n \in D_\Diamond^n$ fail to stand in the relation expressed by P iff some member of $v^-(P)(d_1, \ldots, d_n)$ is a member of $S_@$. Thus, intuitively, we would want to assign the predicate "x is blue" the function that maps every object $d \in D_\Diamond$ to all the possible set of states of the object is being colored other than red: the state of d being blue, the state of d being green, and so on.

The conditions (a), (b), and (c) together ensure that the interpretation that we've just described behaves according to the laws of negative free logic: for every n-ary predicate symbol $P \in \mathcal{P}$, according to condition (c) there can only be an actual state in $v^+(P)(d_1,\dots,d_n)$ if $d_1,\dots,d_n$ are all actual objects, and according to conditions (a) and (b) for all actual objects $d_1,\dots,d_n$ either there is an actual state in $v^+(P)(d_1,\dots,d_n)$ or in $v^-(P)(d_1,\dots,d_n)$—but never in both. Thus, once we've extend the interpretation to *sentences*, these conditions ensure that the logic will be a negative free logic: only statements about existing objects can be true, and there will be neither truth-value *gaps* nor truth-value *gluts*.

3.5.2 Verifiers and Falsifiers of Sentences

We'll now first define the exact verifiers and exact falsifiers for sentences *without the ground operator* under an interpretation in a state space. The clauses we'll use for this are simply a generalization of the clauses given by Fine [42, p. 71–74] to our languages of ground:

Definition 3.5.5. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. We define the sets $[\![\varphi]\!]^+$ of* exact verifiers *and $[\![\varphi]\!]^-$ of* exact falsifiers *for a sentence $\varphi \in \mathcal{L}_{\setminus\leq}$ under $\mathcal{I}$ in $\mathcal{S}$ by simultaneous recursion on the construction of φ:*

(i) *(a)* $[\![c_1 = c_2]\!]^+ = \begin{cases} \{\lambda\} & \textit{if } \delta(c_1) \textit{ and } \delta(c_2) \in D_@ \textit{ and } \delta(c_1) = \delta(c_2) \\ \emptyset & \textit{otherwise} \end{cases}$

(b) $[\![c_1 = c_2]\!]^- = \begin{cases} \{\lambda\} & \textit{if } (\delta(c_1) \notin D_@ \textit{ and } \delta(c_2) \notin D_@) \textit{ or } \delta(c_1) \neq \delta(c_2) \\ \emptyset & \textit{otherwise} \end{cases}$

(ii) *(a)* $[\![P(c_1,\dots,c_n)]\!]^+ = \begin{cases} v^+(P)(\delta(c_1),\dots,\delta(c_n)) & \textit{if } \delta(c_1),\dots,\delta(c_n) \in D_@ \\ \emptyset & \textit{otherwise} \end{cases}$

(b) $[\![P(c_1,\dots,c_n)]\!]^- = \begin{cases} v^-(P)(\delta(c_1),\dots,\delta(c_n)) & \textit{if } \delta(c_1),\dots,\delta(c_n) \in D_@ \\ \{\lambda\} & \textit{otherwise} \end{cases}$

(iii) *(a)* $[\![\neg\varphi]\!]^+ = [\![\varphi]\!]^-$

(b) $[\![\neg\varphi]\!]^- = [\![\varphi]\!]^+$

(iv) *(a)* $[\![\bigwedge\{\varphi_i \mid i \in I\}]\!]^+ = \{\prod\{x_i \mid i \in I\} \mid x_i \in [\![\varphi_i]\!]^+ \textit{ for all } i \in I\}$

(b) $[\![\bigwedge\Gamma]\!]^- = \bigcup\{[\![\varphi]\!]^- \mid \varphi \in \Gamma\} \cup \{\prod[\![\bigvee\Delta]\!]^- \mid \Delta \subseteq \Gamma\}$

(v) *(a)* $[\![\bigvee\Gamma]\!]^+ = \bigcup\{[\![\varphi]\!]^+ \mid \varphi \in \Gamma\} \cup \{[\![\prod\bigwedge\Delta]\!]^+ \mid \Delta \subseteq \Gamma\}$

(b) $[\![\bigvee\{\varphi_i \mid i \in I\}]\!]^- = \{\prod\{x_i \mid i \in I\} \mid x_i \in [\![\varphi_i]\!]^- \textit{ for all } i \in I\}$

These clauses tell us the verifiers and falsifiers for all formulas of a language of ground which don't contain the ground operator $\leq$. Providing clauses for formulas with this operator is the main contribution of this chapter. But before we shall do so, it might be helpful to say a few things about the clauses that we've given here.

As we've said, the clauses of Definition 3.5.5 are essentially the ones of Fine [42, p. 71–74]. But we've generalized the clauses in three ways. First, consider the clauses (i.a–b). Fine doesn't give clauses for identities, simply because he doesn't work with languages that contain the identity symbol. The clauses that we've implemented here satisfy two intuitive constraints. First, the clauses correspond to the idea of negative free logic. To see this, note that the set of exact verifiers of an equation of the form $c_1 = c_2$ is non-empty iff both the denotation $\delta(c_1)$ of c_1 and the denotation $\delta(c_2)$ of c_2 are actual objects and the two denotations are in fact numerically identical. This is exactly the case iff $c_1 = c_2$ would be true under the clauses of negative free logic [104, §3.1]. And conversely, the set of exact falsifiers of an equation of the form $c_1 = c_2$ is non-empty iff the set of its exact verifiers is empty.[13] Thus, for every equation either the set of exact verifiers or the set of exact falsifiers is non-empty. Together with the second constraint, this will effectively give us *bivalence* for equations.

The second intuitive constraint is that we want the truth and falsity of equations to be *fundamental* in a sense. Note that if the set of exact verifiers verifiers is non-empty, then its only member is the null fact λ and analogously for the set of its exact falsifiers. Remember that we've said above that Fine [42, p. 48] proposed that we say that a fact is *zero-grounded* iff it is grounded in the zero fact, and that we can take the zero fact to be λ. Indeed, Fine [42, p. 48] suggests that identities are zero-grounded and in this sense fundamental. Here we reflect this idea by saying that the only exact verifier of a true equation (according to negative free logic) is the zero fact λ and that conversely also the only exact falsifier of a false equation (according to negative free logic) is the zero fact λ. Remember that above we noted that λ is always a member of $S_{@}$. Thus, if we say that an equation is true iff at least one of its exact verifiers is an actual state, we get that an equation is true under our semantics iff it is true under the clauses of negative free logic. And similarly, if we say that an equation is false iff at least one of its exact falsifiers is an actual state, then an equation is false iff is false under the clauses of negative free logic. Moreover, by our previous observation, it follows that equations under this semantics are bivalent.

Now, for the second way in which we've generalized Fine's framework, consider clauses (ii.a–b). Fine only gives clauses for propositional languages and thus we needed to introduce new clauses for atomic sentences of the form $P(c_1, \ldots, c_n)$. As we've explained above, the idea of v^+ and v^- is that they give us the exact conditions for objects to stand in the relation expressed

[13] This follows easily by a single application of the De Morgan laws to the conditions for the sets being non-empty.

by predicates. And here we've simply applied this idea for existing objects: According to clause (ii.a), the set of exact verifiers of a sentence of the form $P(c_1, \ldots, c_n)$ is just the set of exact conditions for the objects denoted by $c_1, \ldots, c_n$ according to δ to stand in the relation expressed by P according to v^+, given that the denotations of $c_1, \ldots, c_n$ are all actual objects. And analogously, clause (ii.b) says that the set of exact falsifiers of a sentence of the form $P(c_1, \ldots, c_n)$ is just the set of exact conditions for the objects denoted by $c_1, \ldots, c_n$ according to δ *not* to stand in the relation expressed by P according to v^-, if all the denotations of $c_1, \ldots, c_n$ are actual objects. If the denotations of $c_1, \ldots, c_n$ are *not* actual objects, then there are no exact verifiers of $P(c_1, \ldots, c_n)$, indeed the zero fact λ is an exact falsifier of the sentence—the falsehood of sentences with constants that don't denote existing objects is fundamental, just like in the case of equations.

The third way in which we've generalized Fine's framework is with regards to clauses (iii.a–b) and (iv.a–b). Fine [42, p. 71–74] only gives the clauses for binary conjunctions and disjunctions. In our clauses, we've generalized Fine's clauses to the corresponding infinitary operations. But the idea stayed the same: the exact verifiers of a conjunction are all the fusion of the exact verifiers of its conjuncts, and the exact falsifiers of a conjunction are all the exact falsifiers of all its conjuncts or arbitrary fusions of some of the exact falsifiers of its conjuncts; and the exact verifiers and falsifiers of a disjunction are simply the dualized exact verifiers and falsifiers of the conjunction of its disjuncts: the exact verifiers of a disjunction are all the exact verifiers of its disjuncts and arbitrary fusions of the exact verifiers of some of its disjuncts, and the exact falsifiers of a disjunction are all the fusions of exact falsifiers of all of its disjuncts. Fine realizes this idea for binary conjunction and disjunction, and we simply generalize the idea to infinitary conjunctions and disjunctions.

Note that according to clause (iv.a) of Definition 3.5.5, $[\![\bigwedge \emptyset]\!]^+ = \{\lambda\}$. In other words, the only exact verifier of the empty conjunction is the zero-fact—what it says is fundamental. This is intuitively plausible since a conjunction of the form $\bigwedge \Gamma$ is intuitively true iff all of its members are. But $\emptyset$ has no members under all circumstances and thus $\bigwedge \emptyset$ is intuitively necessarily true. Indeed the necessary truth of $\bigwedge \emptyset$ appears to be something that is fundamental about our intuitions about conjunctions and thus it's plausible to say that $\bigwedge \emptyset$ is exactly verified by the zero-fact. On the other hand, according to clause (iv.b) Definition 3.5.5, $[\![\bigwedge \emptyset]\!]^- = \emptyset$. In other words, no state exactly falsifies $\bigwedge \emptyset$. By what we just said, also this makes intuitive sense. The empty disjunction $\bigvee \emptyset$ behaves dually: by clause (v.a-b) of Definition 3.5.5, we get that both $[\![\bigwedge \emptyset]\!]^+ = \emptyset$ and $[\![\bigwedge \emptyset]\!]^- = \{\lambda\}$. By dual considerations to the case of the empty conjunctions, these facts are similarly intuitively plausible.

Before we define *truth* and *falsity* under an interpretation, let's make a brief remark about the quantifiers. Remember that we've omitted the quantifiers from our languages of ground. The reason for this omission is that the quantifiers put us even deeper in the realm of infinitary logic. Let's assume for a moment that we the have quantifiers at our disposal. Consider the universal quantifier $\forall$. How could we change our semantics to accommodate $\forall$? We'd have to say what are the exact verifiers and falsifiers of a sentence of the form $\forall x\varphi(x)$ under an interpretation. An intuitively appealing proposal, in fact due to Fine [42, p. 59–63], would be to say that the verifiers of $\forall x\varphi(x)$ are all the fusions of verifiers of $\varphi(c_1), \varphi(c_2), \ldots$ together with the fact that the denotations of of $c_1, c_2, \ldots$ make up all the actual objects, i.e. the fact that $\{\delta(c_1), \delta(c_2), \ldots\} = D_@$. Now the question is how should we express the fact that the denotations of a sequence of terms exhaust the actual objects in the object language? A first approach might be to (syntactically) define a *totality predicate* T by saying that:

$$T(c_i)_{i \in I} =_{def.} \forall x \bigvee \{x = c_i \mid \text{for some } i \in I\}.$$

Then we might think we could say that an exact verifier is the fusion of all the verifiers of $T(c_i)_{i \in I}$ and all the verifiers of all $\varphi(c_i)$ for $i \in I$. But this doesn't really wouldn't get us any further as the definition of T involved the universal quantifier itself: thus in order for this proposal to work, we'd already need to know what the exact verifiers of a universally quantified statement are. For exactly this reason, Fine [42, p. 59–63] proposes to add a *primitive* totality predicate T to our language. But this predicate would have to be a *multigrade predicate* in the sense of Oliver and Smiley [105]. Such predicates are not in themselves problematic, but introducing them would complicate our framework to the point that it would become barely intelligible. For this reason, we postpone treating the full logic of ground with the quantifiers to another time.

Now let us define the notion of truth under an interpretation for sentences without the ground operator. To obtain this definition, we simply formalize the intuitive idea which we've been using all along that a sentence is true iff at least one of its exact verifiers is an actually obtaining state:

Definition 3.5.6. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. We say for all $\varphi \in \mathcal{L}_{\setminus\leq}$ that:*

$$(\mathcal{S}, \mathcal{I}) \vDash \varphi \textit{ iff there is an } x \in [\![\varphi]\!]^+ \textit{ such that } x \in S_@.$$

Given this definition and what we've said to motivate our semantics, it is now easy to see that this notion of truth under an interpretation for sentences without the ground operator behaves like truth under the semantics of (infinitary quantifier free) negative free logic:

Lemma 3.5.7. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. Then,*

(i) $(\mathcal{S}, \mathcal{I}) \vDash P(c_1, \ldots, c_n)$, *only if* $\delta(c_1), \ldots, \delta(c_n) \in D_@$*;*

(ii) $(\mathcal{S}, \mathcal{I}) \vDash c_1 = c_2$ *iff* $\delta(c_1) = \delta(c_2) \in D_@$*;*

(iii) $(\mathcal{S}, \mathcal{I}) \vDash \neg\varphi$ *iff it's not the case that* $(\mathcal{S}, \mathcal{I}) \vDash \varphi$*;*

(iv) $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$ *iff* $(\mathcal{S}, \mathcal{I}) \vDash \varphi$ *for all* $\varphi \in \Gamma$*; and*

(v) $(\mathcal{S}, \mathcal{I}) \vDash \bigvee \Gamma$ *iff* $(\mathcal{S}, \mathcal{I}) \vDash \varphi$ *for some* $\varphi \in \Gamma$*.*

Proof. By inspection of our semantics and our running observations: *(i)* follows by the conditions (a–c) of Definition 3.5.4 and Definition 3.5.6. The other cases, follow from the properties stipulated for $\prod$ in Definition 3.5.1 together with Definition 3.5.4.

□

Indeed, we can even show more. For a given interpretation in a state space, we can define a canonical associated interpretation of (infinitary quantifier free) negative free logic in the traditional sense, under which exactly the same sentences (without the ground operator) are true as in under interpretation in the state space.

Definition 3.5.8. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. We define the associated traditional interpretation $\mathfrak{I} = (D, D_0, d, v)$, where D is the outer domain, D_0 is the inner domain, and d assigns denotations from the outer domain to the constants, and v assigns extensions to the predicates, by setting:*

(i) $D = D_\Diamond$,

(ii) $D_0 = D_@$,

(iii) $d = \delta$,

(iv) $v(P) = \{(d_1, \ldots, d_n) \in D_0^n \mid$ *for some* $x \in S_@, x \in v^+(P)(d_1, \ldots, d_n)\}$.

Then we can show the following lemma:

Lemma 3.5.9. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. Furthermore, let $\mathfrak{I} = (D, D_0, d, v)$ be the associated traditional interpretation from Definition 3.5.8. Then for all $\varphi \in \mathcal{L}_{\setminus\leq}$:*

$$(\mathcal{S}, \mathcal{I}) \vDash \varphi \textit{ iff } \mathfrak{I} \vDash \varphi,$$

where $\mathfrak{I} \vDash \varphi$ is defined in a straightforward way by extending the basic satisfaction clauses of negative free logic with the satisfaction clauses of infinitary propositional logic.

Proof. The only interesting case is for a sentence of the form $P(c_1, \ldots, c_n)$, since the other cases follow by Lemma 3.5.7. Remember that on the basic satisfaction clauses of negative free logic, $\mathfrak{I} \vDash P(c_1, \ldots, c_n)$ iff $d(c_1), \ldots, d(c_n) \in D_0$ and $(d(c_1), \ldots, d(c_n)) \in v(P)$. We show both directions of the claim that $(\mathcal{S}, \mathcal{I}) \vDash P(c_1, \ldots, c_n)$ iff $\mathfrak{I} \vDash P(c_1, \ldots, c_n)$ in turn:

- Assume that $(\mathcal{S}, \mathcal{I}) \vDash P(c_1, \ldots, c_n)$. By Lemma 3.5.7, we already know that $(\mathcal{S}, \mathcal{I}) \vDash P(c_1, \ldots, c_n)$, only if $\delta(c_1) = d(c_1), \ldots, \delta(c_n) = d(c_n) \in D_@ = D_0$. By Definition 3.5.6, we know that $(\mathcal{S}, \mathcal{I}) \vDash P(c_1, \ldots, c_n)$ iff there is some state $x \in S_@$ such that $x \in v^+(P)(\delta(c_1), \ldots, \delta(c_n))$. But by Definition 3.5.8 this is the case iff $\delta(c_1) = d(c_1), \ldots, \delta(c_n) = d(c_n) \in v(P)$.

- For the converse direction, assume that $(\mathcal{S}, \mathcal{I}) \vDash P(c_1, \ldots, c_n)$. By the basic satisfaction clauses of negative free logic, we know that this is the case iff $d(c_1), \ldots, d(c_n) \in D_0$ and $(d(c_1), \ldots, d(c_n)) \in v(P)$. By Definition 3.5.8, this is the case iff there is some $x \in S_@$ such that $x \in v^+(P)(d(c_1), \ldots, d(c_n))$. Since $\delta(c_1) = d(c_1), \ldots, \delta(c_n) = d(c_n)$, the claim follows.

$\square$

This lemma gives us a precise sense in which our semantics from this section is just a more fine-grained version of the traditional semantics for (infinitary quantifier free) negative free logic. Intuitively, by means of the exact verifiers and falsifiers of sentences, the semantics does not only tell us *whether* a sentence is true or false, but also *why* the sentence is true. So far, however, we haven't made use of this additional information. In the next section, we'll use this additional information to give semantic clauses for sentences with the ground operator in them.

3.5.3 Clauses for the Ground Operator

Fine [42, 44] gives semantic clauses for the *truth* of a statement of ground of the form $\Gamma \leq \varphi$, where both $\Gamma \subseteq \mathcal{L}_{\backslash\leq}$ and $\varphi \in \mathcal{L}_{\backslash\leq}$ are sentences without the ground operator. Fine's idea is that we can understand weak ground as transmission of exact verifiers: we can say that the fact expressed by a (true) sentence $\varphi \in \mathcal{L}_{\backslash\leq}$ holds in virtue of the facts expressed by the sentences in a set $\Gamma \subseteq \mathcal{L}_{\backslash\leq}$ iff all the sentences in Γ are true and all the fusions of any sequence of exact verifiers for all the members of Γ are also exact verifiers of φ. Thus, for the fact that φ to hold in virtue of the facts expressed by the members of Γ is for all the members of Γ to be true and, intuitively, for the reasons why the members of Γ are true together to be a reason for φ to be true. This gives us a very intuitive notion of one fact holding in virtue of others on the weak concept of ground.

By the interdefinability of strict and weak ground, Fine's clause also gives us truth conditions for strict statements of ground of the form $\Gamma < \varphi$, where both $\Gamma \subseteq \mathcal{L}_{\setminus\leq}$ and $\varphi \in \mathcal{L}_{\setminus\leq}$ are sentences without the ground operator. By simply applying the definition of strict ground in terms of weak ground, we get that $\Gamma < \varphi$ is true iff $\Gamma \leq \varphi$ is true, but for no $\psi \in \Gamma$ is there a set of sentences Δ such that $\varphi, \Delta \leq \psi$ is true. In other words, $\Gamma < \varphi$ is true iff all the sentences in Γ are true and all the fusions of any sequence of exact verifiers for all the members of Γ are also exact verifiers of φ and no exact verifier of φ can be fused with some other states such that we get an exact verifier of some member of Γ. This gives us a very intuitive notion of one fact holding in virtue of others on the strict concept of ground.

Let us turn these informal considerations into a precise definition:

Definition 3.5.10. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. Then we say for all $\{x_i \mid i \in I\} \subseteq \mathcal{L}_{<\lambda}$ and $\psi \in \mathcal{L}_{\setminus\leq}$:*

- $(\mathcal{S}, \mathcal{I}) \vDash \{\varphi_i \mid i \in I\} \leq \psi$ *iff*
  - *(i) for some family $(x_i \in S_\Diamond)_{i \in I}$ of states such that $x_i \in [\![\varphi_i]\!]^+$ for all $i \in I$, we have $\prod\{x_i \mid i \in i\} \in S_@$,*
  - *(ii) for all families $(x_i \in S_\Diamond)_{i \in I}$ such that $x_i \in [\![\varphi_i]\!]^+$ for all $i \in I$, then $\prod\{x_i \mid i \in I\} \in [\![\psi]\!]^+$.*

This definition is quite a mouthful. But using clause (iv.a) of Definition 3.5.5 the definition can be significantly simplified. Using this clause we can equivalently restate by saying for all languages of ground $\mathcal{L}$, all state spaces $\mathcal{S}$, and all interpretations $\mathcal{I}$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$,

- $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq \psi$, where $\Gamma \subseteq \mathcal{L}_{\setminus\leq}$ and $\psi \in \mathcal{L}_{\setminus\leq}$, iff
  - (i) $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$, and
  - (ii) $[\![\bigwedge \Gamma]\!]^+ \subseteq [\![\varphi]\!]^+$.

Note that, according to this definition, we get for a sentence of the form $\emptyset \leq \varphi$, for $\varphi \in \mathcal{L}$, that $(\mathcal{S}, \mathcal{I}) \vDash \emptyset \leq \varphi$ iff $\lambda \in [\![\varphi]\!]^+$. Since the zero-state λ is always part of the actual states, we can use sentences of the form $\emptyset \leq \varphi$ to express that the fact expressed by φ is fundamental (in Fine's sense of zero-grounded): the zero-state is already directly responsible for the truth of φ.

It is easy to see now, that this definition validates the intuitive principles for weak ground, such as **Reflexivity**, **Transitivity**, and the factivity laws:

Lemma 3.5.11. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S}$ a state space, and $\mathcal{I}$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. Then for all $\Gamma \subseteq \mathcal{L}_{\setminus\leq}$ and $\psi \in \mathcal{L}_{\setminus\leq}$,*

(i) if $(\mathcal{S}, \mathcal{I}) \vDash \varphi$, then $(\mathcal{S}, \mathcal{I}) \vDash \varphi \leq \varphi$,

(ii) if $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq \varphi$ and $(\mathcal{S}, \mathcal{I}) \vDash \varphi, \Delta \leq \psi$, then $(\mathcal{S}, \mathcal{I}) \vDash \Gamma, \Delta \leq \psi$,[14]

(iii) if $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq \varphi$, then $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$, and

(iv) if $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq \varphi$, then $(\mathcal{S}, \mathcal{I}) \vDash \varphi$.

Proof. For the first claim (i) assume that $(\mathcal{S}, \mathcal{I}) \vDash \varphi$. Since $(\mathcal{S}, \mathcal{I}) \vDash \varphi$ and $[\![\varphi]\!]^+ \subseteq [\![\varphi]\!]^+$ is trivially true, the claim follows immediately. For second claim (ii), assume that both $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq \varphi$ and $(\mathcal{S}, \mathcal{I}) \vDash \varphi, \Delta \leq \psi$. It follows that $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$ and $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge\{\varphi\} \cup \Delta$, and hence using Lemma 3.5.7 that $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma \cup \Delta$. Next we need to check that $[\![\bigwedge \Gamma \cup \Delta]\!]^+ \subseteq [\![\varphi]\!]^+$. We know by assumption that $[\![\bigwedge \Gamma]\!]^+ \subseteq [\![\psi]\!]^+$ and $[\![\bigwedge\{\varphi\} \cup \Delta]\!]^+ \subseteq [\![\varphi]\!]^+$. But from this, using clause (d) of Definition 3.5.1 and clause (iv.a) of Definition 3.5.5, it follows that $[\![\bigwedge \Gamma \cup \Delta]\!]^+ \subseteq [\![\varphi]\!]^+$. The third claim (iii) follows immediately by clause (i) of Definition 3.5.3. And the fourth claim follows immediately by putting together (iii), clause (ii) of Definition 3.5.3, and Definition 3.5.6. □

Moreover, using this lemma and Definition 3.4.3, we can establish that also the defined notion of strict ground satisfies the desired properties, such as **Irreflexivity**, **Transitivity**, and the factivity laws:

Lemma 3.5.12. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S}$ a state space, and $\mathcal{I}$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. Then for all $\Gamma \subseteq \mathcal{L}_{\setminus\leq}$ and $\psi \in \mathcal{L}_{\setminus\leq}$,*

(i) for no $\Gamma \subseteq \mathcal{L}_{\setminus\leq}$ and $\varphi \in \mathcal{L}_{\setminus\leq}$ we have $(\mathcal{S}, \mathcal{I}) \vDash \varphi, \Gamma < \varphi$,

(ii) if $(\mathcal{S}, \mathcal{I}) \vDash \Gamma < \varphi$ and $(\mathcal{S}, \mathcal{I}) \vDash \varphi, \Delta < \psi$, then $(\mathcal{S}, \mathcal{I}) \vDash \Gamma, \Delta < \psi$,

(iii) if $(\mathcal{S}, \mathcal{I}) \vDash \Gamma < \varphi$, then $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$, and

(iv) if $(\mathcal{S}, \mathcal{I}) \vDash \Gamma < \varphi$, then $(\mathcal{S}, \mathcal{I}) \vDash \varphi$.

Proof. Remember that by Definition 3.4.3, we have that

$$\Gamma < \varphi =_{def} (\Gamma \leq \varphi) \wedge \bigwedge \{\neg(\varphi, \Delta \leq \psi) \mid \Delta \subseteq \mathcal{L}, \psi \in \Gamma\}.$$

[14] There is also an *infinitary* version of this principle, which allows us to "cut" infinitely many formulas at a time [44]. A rule corresponding to this principle is required for completeness proofs in the logic of ground, but since we're not primarily interested in completeness here, for reasons of perspicuity, we'll stick with the more simple principle, which allows us to "cut" only one formula at a time.

Now, for the first claim (i), we can assume without constraint of generality that both $(\mathcal{S}, \mathcal{I}) \vDash \varphi$ and $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$, because otherwise the claim would hold immediately by clause (i) of Definition 3.5.3. But then the first claim (i) follows from the fact (i) of Lemma 3.5.11 gives us that $(\mathcal{S}, \mathcal{I}) \vDash \varphi \leq \varphi$. And from this using the laws of Lemma 3.5.7, we can infer that $(\mathcal{S}, \mathcal{I}) \vDash \neg \bigwedge \{\neg(\varphi, \Delta \leq \varphi) \mid \Delta \subseteq \mathcal{L}\}$, since we may simply take $\Delta = \emptyset$. Hence again by the same laws, $(\mathcal{S}, \mathcal{I}) \vDash \neg((\Gamma \leq \varphi) \wedge \bigwedge \{\neg(\varphi, \Delta \leq \varphi) \mid \Delta \subseteq \mathcal{L}\})$. The claim follows by Definition 3.4.3. For the proof of the second claim we use (ii) of Lemma 3.5.11 and the laws of Lemma 3.5.7 to show that if $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \{\neg(\varphi, \Delta \leq \psi) \mid \Delta \subseteq \mathcal{L}, \psi \in \Gamma\}$ and $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \{\neg(\psi, \Sigma \leq \theta) \mid \Sigma \subseteq \mathcal{L}, \psi \in \Delta \cup \{\varphi\}\}$, then $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \{\neg(\varphi, \Sigma \leq \theta) \mid \Sigma \subseteq \mathcal{L}, \psi \in \Delta \cup \Gamma\}$. Then, the claim follows by putting this and (ii) of Lemma 3.5.11 together. The claims (iii) and (iv) follow immediately from Lemma 3.5.11 and Definition 3.4.3.

□

Finally, using the definition of $\leftrightharpoons$ in Definition 3.4.3, we can show that $\leftrightharpoons$ really captures a concept of factual equivalence:

Lemma 3.5.13. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S}$ a state space, and $\mathcal{I}$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. Then for all $\Gamma \subseteq \mathcal{L}_{\setminus\leq}$*

(i) if $(\mathcal{S}, \mathcal{I}) \vDash \varphi$ and $(\mathcal{S}, \mathcal{I}) \vDash \psi$, then $(\mathcal{S}, \mathcal{I}) \vDash \varphi \leftrightharpoons \psi$ iff $[\![\varphi]\!]^+ = [\![\psi]\!]^+$; and

(ii) if $(\mathcal{S}, \mathcal{I}) \vDash \neg\varphi$ and $(\mathcal{S}, \mathcal{I}) \vDash \neg\psi$, then $(\mathcal{S}, \mathcal{I}) \vDash \neg\varphi \leftrightharpoons \neg\psi$ iff $[\![\varphi]\!]^- = [\![\varphi]\!]^-$.

Proof. Remember that according to Definition 3.4.3, we have that

$$\varphi \leftrightharpoons \psi =_{def} (\varphi \leq \psi) \wedge (\psi \leq \varphi).$$

But then the first claim (i) simply follows by observing that $[\![\varphi]\!]^+ = [\![\varphi]\!]^+$ iff both $[\![\varphi]\!]^+ \subseteq [\![\psi]\!]^+$ and $[\![\psi]\!]^+ \subseteq [\![\varphi]\!]^+$ and the condition (ii) of Definition 3.5.3. And the second claim (ii) follows from (i) by applying clause (iii.a) of Definition 3.5.5 to $\neg\varphi$ and $\neg\psi$. □

This lemma gives us a precise sense in which the defined operator $\leftrightharpoons$ captures the informal notion of factual equivalence that we've sketched on p. 118.[15]

We have seen that Fine's semantic clauses for the truth of a statement of ground of the form $\Gamma \leq \varphi$, where both $\Gamma \subseteq \mathcal{L}_{\setminus\leq}$ and $\varphi \in \mathcal{L}_{\setminus\leq}$ are sentences without the ground operator, behave like we want them to: they

[15] Indeed, it can be shown that this concept satisifes (a factive version) of Angell's system of analytic equivalence. The proof is not very complicated, but we omit it here so as not to detract to much from the main issue. For the details, we refer the reader to [37].

semantically validate the intuitively valid principles of ground. Effectively, they achieve this by using the more fine-grained structure of state spaces under an interpretation to explain, in an intuitive way, what has to be the case for a statement of ground to be true. But the clauses don't embed statements of ground themselves into the fine-grained structure of the state spaces: the clauses only tell us *whether* a statement of ground is true and not *why*. More specifically, they don't give us the exact verifiers and falsifiers of statements of ground.

For this reason, we can't apply Fine's clauses to iterated statements of ground, which may be of the form

$$\Gamma, (\Delta \leq \varphi) \leq \psi \text{ or}$$

$$\Gamma \leq (\Delta \leq \varphi),$$

where $\Gamma, \Delta \subseteq \mathcal{L}$ and $\varphi, \psi \in \mathcal{L}$. For if we wanted to apply the clause from Definition 3.5.3 to these sentences, we'd have to know what the exact verifiers of statements of the form $\Delta \leq \varphi$ or $\Delta \leq \varphi$ are: otherwise we can show neither that $[\![\bigwedge \Gamma \cup \{\Delta \leq \varphi\}]\!]^+ \subseteq [\![\psi]\!]^+$ nor that $[\![\bigwedge \Gamma]\!]^+ \subseteq [\![\Delta \leq \varphi]\!]^+$. So, if we wish to define the truth conditions of iterated statements of ground using something like Fine's clauses, then we need to say what are the exact verifiers (and falsifiers) of statements of ground.

Now, note that if we assign all the fusions of the exact verifiers of some sentences Γ to be the exact verifiers of a statement of the form $\Delta \leq \varphi$, then, following Fine's clauses, we will get that $\Gamma \leq (\Delta \leq \varphi)$ is true. This observation provides us with a heuristic for giving the exact verifiers of statements of ground: we simply take the exact verifiers of a true statement to be the fusions of the exact verifiers of the sentences that we wish to say the fact expressed by the statement of ground holds in virtue of. In §3.3.2 we have discussedsome views about this and we've said that in this paper we'd like to determine the logic of iterated ground according to the view that grounds ground ground. Thus, we get that we should take the exact verifiers of a true statement of the form $\Gamma \leq \varphi$ to be all the fusions of all the verifiers of the members of Γ—in other words, we should take them to be the verifiers of $\bigwedge \Gamma$. In contrast, a *false* statement of the form $\Gamma \leq \varphi$ should not have *any* exact verifiers, only exact falsifiers.

So what should we talk the exact falsifiers of a false statement of ground to be? According to Fine's clauses, we can distinguish two ways in which a statement of the form $\Gamma \leq \varphi$ could be false under an interpretation $\mathcal{I}$ in a state spaces $\mathcal{S}$. First, it could be the case that we don't have $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$, regardless of whether $[\![\bigwedge \Gamma]\!]^+ \subseteq [\![\varphi]\!]^+$ or not; and second we could have $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$, but not $[\![\bigwedge \Gamma]\!]^+ \subseteq [\![\varphi]\!]^+$. In the first case, on the view that grounds ground ground, it seems to be clear why the statement is false:

because at least one of the grounds fails, i.e. $\bigwedge\Gamma$ is false. And thus a good candidate for the exact falsifiers of $\Gamma \leq \varphi$ in this case would be the exact falsifiers of $\bigwedge\Gamma$: i.e. simply $[\![\bigwedge\Gamma]\!]^+$. In the other case, we might be at a loss at first: there is no possible state in our state space that corresponds to the fact that $[\![\bigwedge\Gamma]\!]^+ \not\subseteq [\![\varphi]\!]^+$—if this is the case, it appears to be a fundamental fact about our state space. But this actually suggests a natural candidate for what we could take the exact falsifier of a false statement of the form $\Gamma \leq \varphi$ to be in this case: since the fact that is responsible for its falsehood is a fundamental fact about the state space, we could take the zero-state λ as the exact falsifier in this case.

Putting these observations together, we arrive at the following definition:

Definition 3.5.14. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. We define the sets $[\![\varphi]\!]^+$ of* exact verifiers *and $[\![\varphi]\!]^-$ of* exact falsifiers *for a sentence $\varphi \in \mathcal{L}$ under $\mathcal{I}$ in $\mathcal{S}$ by a simultaneous recursion on the construction of φ, which consists of the clauses (i–v) of Definition 3.5.5 plus:*

(vi) *(a)* $$[\![\Gamma \leq \varphi]\!]^+ = \begin{cases} [\![\bigwedge\Gamma]\!]^+ & \text{if } (\mathcal{S},\mathcal{I}) \vDash \bigwedge\Gamma \text{ and } [\![\bigwedge\Gamma]\!]^+ \subseteq [\![\varphi]\!]^+ \\ \emptyset & \text{otherwise} \end{cases}$$

(b) $$[\![\Gamma \leq \varphi]\!]^- = \begin{cases} [\![\bigwedge\Gamma]\!]^- & \text{if not } (\mathcal{S},\mathcal{I}) \vDash \bigwedge\Gamma \\ \{\lambda\} & \text{if } (\mathcal{S},\mathcal{I}) \vDash \bigwedge\Gamma, \text{ but not } [\![\bigwedge\Gamma]\!]^+ \subseteq [\![\varphi]\!]^+ \\ \emptyset & \text{otherwise} \end{cases}$$

This definition gives us the natural exact verifiers and falsifiers of statements of ground under the view that grounds ground ground.

To determine logic given by the above definition, let's first generalize Definition 3.5.6 to all formulas of our languages of ground:

Definition 3.5.15. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. We say for all $\varphi \in \mathcal{L}$ that:*

$$(\mathcal{S},\mathcal{I}) \vDash \varphi \text{ iff } [\![\varphi]\!]^+ \cap S_@ \neq \emptyset.$$

Then we say for all $\Gamma \subseteq \mathcal{L}$ and $\varphi \in \mathcal{L}$ that $\Gamma \vDash \varphi$ iff for all state spaces $\mathcal{S}$ and all interpretations $\mathcal{I}$ for $\mathcal{L}$ in $\mathcal{S}$, if $(\mathcal{S},\mathcal{I}) \vDash \psi$, for all $\psi \in \Gamma$, then $(\mathcal{S},\mathcal{I}) \vDash \varphi$.

The first thing to note is that Fine's clause for $(\mathcal{S},\mathcal{I}) \vDash \Gamma \leq \varphi$ from Definition 3.5.3 in fact agrees with our clause in Definition 3.5.15:

Lemma 3.5.16. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$ a state space, and $\mathcal{I} = (\delta, v^+, v^-)$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. If we write $(\mathcal{S},\mathcal{I}) \vDash_1 \Gamma \leq \varphi$ for the satisfaction relation defined in Definition 3.5.3 and $(\mathcal{S},\mathcal{I}) \vDash_2 \Gamma \leq \varphi$ for the satisfaction relation defined in Definition 3.5.15, then we get:*

$$(\mathcal{S}, \mathcal{I}) \vDash_1 \Gamma \leq \varphi \textit{ iff } (\mathcal{S}, \mathcal{I}) \vDash_2 \Gamma \leq \varphi.$$

Proof. Left to right: If $(\mathcal{S}, \mathcal{I}) \vDash_1 \Gamma \leq \varphi$, then both $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$ and $[\![\bigwedge \Gamma]\!]^+ \subseteq [\![\varphi]\!]^+$ by definition. But then by clause (vi.a) $[\![\Gamma \leq \varphi]\!]^+ = [\![\bigwedge]\!]^+$. And since $(\mathcal{S}, \mathcal{I}) \vDash_1 \bigwedge \Gamma$, which by Definition 3.5.6 means that there is some $x \in [\![\bigwedge \Gamma]\!]^+$ such that $x \in S_@$, we get that there is some $x \in [\![\Gamma \leq \varphi]\!]^+ = [\![\bigwedge]\!]^+$ such that $x \in S_@$, as desired.

Right to left: If $(\mathcal{S}, \mathcal{I}) \vDash_2 \Gamma \leq \varphi$, then $x \in [\![\Gamma \leq \varphi]\!]^+$ such that $x \in S_@$. But by Definition 3.5.14 $[\![\Gamma \leq \varphi]\!]^+$ is non-empty iff both $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$ and $[\![\bigwedge \Gamma]\!]^+ \subseteq [\![\varphi]\!]^+$. Hence we get by Definition 3.5.3 that $(\mathcal{S}, \mathcal{I}) \vDash_1 \Gamma \leq \varphi$, as desired.

□

Thus, we can drop the subscripts in the following. Moreover, as a consequence of this lemma, the Lemmas 3.5.7, 3.5.11, 3.5.12, and 3.5.13 all hold in full generality, meaning for all sentences of $\mathcal{L}$. In other words, the semantics given by Definitions 3.5.14 and 3.5.15 gives us an intuitively adequate semantics for the concept of weak ground, and, via the definition of strict ground in terms of weak ground, also for the concept of strict ground.

The following lemma shows that Definitions 3.5.14 and 3.5.15 really give us a semantics of iterated ground that corresponds to the view that grounds ground ground:

Lemma 3.5.17. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S}$ a state space, and $\mathcal{I}$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. Then, for all $\Gamma \subseteq \mathcal{L}$ and $\varphi \in \mathcal{L}$,*

(i) (a) if $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq \varphi$, then $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq (\Gamma \leq \varphi)$;

(b) if $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq (\Delta \leq \varphi)$, then $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq \bigwedge \Delta$;

(ii) if $(\mathcal{S}, \mathcal{I}) \vDash \neg \bigwedge \Gamma$, then $(\mathcal{S}, \mathcal{I}) \vDash (\neg \bigwedge \Gamma) \leq \neg(\Gamma \leq \varphi)$,

(iii) if $(\mathcal{S}, \mathcal{I}) \vDash \bigwedge \Gamma$ and $(\mathcal{S}, \mathcal{I}) \vDash \neg(\Gamma \leq \varphi)$, then $(\mathcal{S}, \mathcal{I}) \vDash \emptyset \leq \neg(\Gamma \leq \varphi)$.

Proof. Immediate by inspection of the truth-conditions for $\Gamma \leq \varphi$. The claims (i.a-b) follow immediately from the fact that if $(\mathcal{S}, \mathcal{I}) \vDash \Gamma \leq \varphi$, then $[\![\Gamma \leq \varphi]\!]^+ = [\![\bigwedge \Gamma]\!]^+$. The claims (ii) and (iii) follow by an inspection of the conditions of clause (vi.b) of Definition 3.5.14. □

This lemma gives us a precise sense in which our semantics captures the view grounds ground ground: in particular, (i.a) corresponds directly to the characteristic inference from (1) to (3) on p. 110, §3.3.2; the clauses (ii-iii), on the other hand, we can take as "rounding things out" in the negative that we didn't talk about in §3.3.2.

Finally, we'd like to know how weak ground and the logical vocabulary interact on the semantics that we have given in this section. The answer is given by the following lemma:

Lemma 3.5.18. *Let $\mathcal{L}$ be a language of ground, $\mathcal{S}$ a state space, and $\mathcal{I}$ and interpretation for $\mathcal{L}$ in $\mathcal{S}$. Then,*

(i) if $(\mathcal{S},\mathcal{I}) \vDash c = c$, then $(\mathcal{S},\mathcal{I}) \vDash \emptyset \leq c = c$;

(ii) if $(\mathcal{S},\mathcal{I}) \vDash \neg(c = c)$, then $(\mathcal{S},\mathcal{I}) \vDash \emptyset \leq \neg(c = c)$

(iii) if $(\mathcal{S},\mathcal{I}) \vDash \varphi_i$, for some $i \in I$, then $(\mathcal{S},\mathcal{I}) \vDash \varphi_i \leq \bigvee_{i\in I} \varphi_i$;

(iv) if $(\mathcal{S},\mathcal{I}) \vDash \varphi_i$, for all $i \in I$, then $(\mathcal{S},\mathcal{I}) \vDash \{\varphi_i \mid i \in I\} \leq \bigwedge_{i\in I} \varphi_i$;

(v) if $(\mathcal{S},\mathcal{I}) \vDash \Gamma, \Delta \leq \varphi$, then $(\mathcal{S},\mathcal{I}) \vDash \Gamma, \bigwedge \Delta \leq \varphi$, given that $\Delta \neq \emptyset$;

(vi) if $(\mathcal{S},\mathcal{I}) \vDash \Gamma, \Delta \leq \varphi$, then $(\mathcal{S},\mathcal{I}) \vDash \Gamma, \bigvee \Sigma \leq \varphi$, given that $\Sigma \subseteq \Delta$ and $\Sigma \neq \emptyset$;

(vii) if $(\mathcal{S},\mathcal{I}) \vDash \varphi$, then $(\mathcal{S},\mathcal{I}) \vDash \varphi \leftrightharpoons \neg\neg\varphi$;

(viii) if $(\mathcal{S},\mathcal{I}) \vDash \neg \bigvee_{i\in I} \varphi_i$, then $(\mathcal{S},\mathcal{I}) \vDash \neg \bigvee_{i\in I} \varphi_i \leftrightharpoons \bigwedge_{i\in I} \neg\varphi_i$;

(ix) if $(\mathcal{S},\mathcal{I}) \vDash \neg \bigwedge_{i\in I} \varphi_i$, then $(\mathcal{S},\mathcal{I}) \vDash \neg \bigwedge_{i\in I} \varphi_i \leftrightharpoons \bigvee_{i\in I} \neg\varphi_i$; and

(x) if $(\mathcal{S},\mathcal{I}) \vDash \bigwedge_{i\in I} \bigvee_{j\in J_i} \varphi_{i,j}$, for some double indexing of formulas $\varphi_{i,j}$ with $i \in I$ and $j \in J_i$, then

$$(\mathcal{S},\mathcal{I}) \vDash \bigwedge_{i\in I} \bigvee_{j\in J_i} \varphi_{i,j} \leftrightharpoons \bigvee_{f\in F} \bigwedge_{i\in I} \varphi_{i,f(i)},$$

where F is a set of choice functions *that chose for each $i \in I$ an index $f(i) \in J_i$.*

Proof. The theorem is a generalization of the main results in [37] and the proof idea is the same. However, we will not give the proof here as many cases rely on relatively strong infinitary principles. For example, showing the infinitary distributivity law (viii) requires (repeated) applications of the axiom of choice. But for illustration, let us prove (i) and (vii). For (i) assume that $(\mathcal{S},\mathcal{I}) \vDash c = c$, meaning there is an $x \in [\![c = c]\!]^+$ such that $x \in S_@$. We've already observed that $[\![c = c]\!]^+$ is non-empty iff $[\![c = c]\!]^+ = \{\lambda\}$. But we've already observed that for all φ, $(\mathcal{S},\mathcal{I}) \vDash \emptyset \leq \varphi$ iff $\lambda \in [\![\varphi]\!]^+$. Hence the claim follows. For (vii) first note that by clause (iii) of Definition 3.5.5, for all φ, $[\![\neg\varphi]\!]^+ = [\![\varphi]\!]^-$ and $[\![\neg\varphi]\!]^- = [\![\varphi]\!]^+$. Thus, by a double application of this, $[\![\neg\neg\varphi]\!]^+ = [\![\varphi]\!]^+$ and $[\![\neg\neg\varphi]\!]^- = [\![\varphi]\!]^-$. Now the claim follows immediately by Lemma 3.5.13.

□

3.6 The Full Logic of Ground

In the last section, we've presented a semantics for the ground operator that can accommodate iterated statements of ground. A natural question now is: What is the logic of (iterated) ground determined by this semantics? Lemmas 3.5.7, 3.5.11, 3.5.12, 3.5.13, and 3.5.13 already give us a pretty good idea of the logic. But how should we axiomatize it? In this section, we'll propose an infinitary logic for this purpose. The previously mentioned lemmas will give us a heuristic for the rules that we should use and indeed they will immediately establish soundness. However, completeness is a different issue. As we've mentioned in §3.4.4, showing a completeness result requires quite substantial infinitary methods, and for this reason we'll omit the result. In the conclusion of this chapter, however, we'll discuss how a completeness result may be obtained.

3.6.1 Natural Deduction for the Full Logic of Ground

We propose to use a *natural deduction system* for the full logic of ground. To obtain such a system, we first need a background logic. For this purpose, we simply extend the standard natural deduction rules for infinitary logic with identity rules that correspond to the rules of negative free logic. In particular, we'll assume the following rules inspired by Lemma 3.5.7:

$$\frac{\varphi(c)}{c = c} \qquad \frac{c_1 = c_2 \quad \varphi(c_1)}{\varphi(c_2)} \qquad \frac{\begin{matrix}[\varphi]\\ \vdots \\ \bot\end{matrix}}{\neg\varphi} \qquad \frac{\neg\varphi \quad \varphi}{\bot}$$

$$\frac{(\varphi_i)_{i\in I}}{\bigwedge_{i\in I}\varphi_i} \qquad \frac{\bigvee_{i\in I}\varphi_i \quad \begin{matrix}[(\varphi_i)_{i\in I}]\\ \vdots \\ \theta\end{matrix}}{\theta} \qquad \frac{\varphi_i}{\bigvee_{i\in I}\varphi_i} \qquad \frac{\bigvee_{i\in I}\varphi_i \quad \begin{matrix}[\varphi_i]\\ \vdots\ \text{for all } i \\ \theta\end{matrix}}{\theta}$$

$$\frac{\begin{matrix}[\varphi]\\ \vdots \\ \theta\end{matrix} \quad \begin{matrix}[\neg\varphi]\\ \vdots \\ \theta\end{matrix}}{\theta} \qquad \frac{\bot}{\varphi}$$

A few remarks about these rules: The first identity rule corresponds to the standard assumption of negative free logic that a formula can only be true if all the constants in it denote existing objects. We've already observed that a constant c denotes an actual object under an interpretation iff $c = c$ is true according to our semantics. Thus, the first rule captures the assumption of negative free logic. The second identity rule is the standard rule of the substitutivity of identicals, which is obviously sound under our semantics. The other rules are merely natural deduction rules for infinitary propositional logic. Note, however, that the resulting proof system will, of course, be infinitary: rules like the introduction rule for $\bigwedge$ may have arbitrarily many premises. Derivations can, nevertheless, only be infinitely *wide* and never infinitely *long*. These rules are all sound by Lemma 3.5.7: they always lead from true premises to true conclusions.

To capture the interaction of the logical vocabulary and the ground operator, we propose the following rules inspired by Lemmas 3.5.11, 3.5.17, and 3.5.18:

$$
\begin{array}{cccc}
\dfrac{\varphi}{\varphi \leq \varphi} & \dfrac{\Gamma \leq \psi \quad \psi, \Delta \leq \varphi}{\Gamma, \Delta \leq \varphi} & \dfrac{\{\varphi_i \mid i \in I\} \leq \psi}{\varphi_i} & \dfrac{\{\varphi_i \mid i \in I\} \leq \psi}{\psi} \\[2ex]
\dfrac{\Gamma \leq \varphi}{\Gamma \leq (\Gamma \leq \varphi)} & \dfrac{(\Gamma \leq (\Delta \leq \varphi)}{\Gamma \leq \bigwedge \Delta} & \dfrac{\neg \bigwedge \Gamma}{(\neg \bigwedge \Gamma) \leq \neg(\Gamma \leq \varphi)} & \dfrac{\bigwedge \Gamma \quad \neg(\Gamma \leq \varphi)}{\emptyset \leq \neg(\Gamma \leq \varphi)} \\[2ex]
\dfrac{\varphi_i}{\varphi_i \leq \bigvee_{i \in I} \varphi_i} & \dfrac{(\varphi_i)_{i \in I}}{\{\varphi_i \mid i \in I\} \leq \bigwedge_{i \in I} \varphi_i} & \dfrac{\Gamma, \Delta \leq \varphi}{\Gamma, \bigwedge \Delta \leq \varphi}\ \Delta \neq \emptyset & \dfrac{\Gamma, \Delta \leq \varphi}{\Gamma, \bigvee \Sigma \leq \varphi}\ \emptyset \neq \Delta \subseteq \Sigma \\[2ex]
\dfrac{\varphi}{\varphi \leftrightharpoons \neg\neg\varphi} & \dfrac{\neg \bigvee_{i \in I} \varphi_i}{\neg \bigvee_{i \in I} \varphi_i \leftrightharpoons \bigwedge_{i \in I} \neg\varphi_i} & \dfrac{\neg \bigwedge_{i \in I} \varphi_i}{\neg \bigwedge_{i \in I} \varphi_i \leftrightharpoons \bigvee_{i \in I} \neg\varphi_i} & \dfrac{c_1 = c_2}{\emptyset \leq c_1 = c_2}
\end{array}
$$

$$
\dfrac{\neg(c_1 = c_2)}{\emptyset \leq \neg(c_1 = c_2)} \qquad \dfrac{\bigvee_{f \in F} \bigwedge_{i \in I} \varphi_{i,f(i)}}{\bigwedge_{i \in I} \bigvee_{j \in J_i} \varphi_{i,j} \leftrightharpoons \bigvee_{f \in F} \bigwedge_{i \in I} \varphi_{i,f(i)}}\ \dagger
$$

†: For some double indexing of formulas $\varphi_{i,j}$ with $i \in I$ and $j \in J_i$, and F a set of choice functions that chose for each $i \in I$ and index $f(i) \in J_i$.

By Lemmas 3.5.11, 3.5.17 and 3.5.18 it follows immediately that these rules are also sound: they always lead from true premises to true conclusions.

Let's call the logic that consists of all the above rules (over some language of ground) the *negative full logic of ground* $NFLG$ and let's denote the (standardly defined) derivability relation in $NFLG$ by $\vdash_{NFLG}$. We do not wish to venture far into the proof theory of $NFLG$, but let's remark a couple of facts. First, as we've already mentioned, $NFLG$ is sound for our semantics:

Proposition 3.6.1. *Let $\mathcal{L}$ a language of ground. Then $NFLG$ over $\mathcal{L}$ is sound with regard to the semantics we defined in the last section. In particular, for all $\Gamma \subseteq \mathcal{L}$ and $\varphi \in \mathcal{L}$:*

$$\textit{if } \Gamma \vdash \varphi, \textit{ then } \Gamma \vDash \varphi.$$

Proof. As we've already observed, this follows immediately by Lemmas 3.5.7, 3.5.11, 3.5.17, and 3.5.18. □

And second, the logic that we've describe here proves (a version of) what Fine calls the *pure logic of full ground* $PLFG$ (not to be confused with what we call the full logic of ground) which consists of the following rules (over our background language $\mathcal{L}$):

$$\frac{\varphi}{\varphi \leq \varphi} \quad \frac{\Gamma < \varphi}{\Gamma \leq \varphi} \quad \frac{\Gamma \leq \psi \quad \psi, \Delta \leq \varphi}{\Gamma, \Delta \leq \varphi} \quad \frac{\Gamma \leq \psi \quad \psi, \Delta < \varphi}{\Gamma, \Delta < \varphi} \quad \frac{\varphi, \Gamma < \varphi}{\bot}$$

$$\frac{\varphi_1, \varphi_2, \ldots \leq \psi \quad \varphi_1, \Gamma_1 < \theta_1 \quad \theta_1, \Delta_1 \leq \varphi \quad \varphi_2, \Gamma_2 < \theta_2 \quad \theta_2, \Delta_2 \leq \varphi \quad \ldots}{\varphi_1, \varphi_2, \ldots < \psi}$$

Let's denote the derivability relation of this logic by $\vdash_{PLFG}$. Then we can show the following lemma:

Lemma 3.6.2. *Let $\mathcal{L}$ be a logic of ground. Then for all $\Gamma \subseteq \mathcal{L}$ and $\varphi \in \mathcal{L}$, if $\Gamma \vdash_{PLFG} \varphi$, then $\Gamma \vdash_{NFLG} \varphi$.*

Proof. The proof is pretty straightforward using the definition of strict ground in terms of weak ground

$$\Gamma < \varphi =_{def} (\Gamma \leq \varphi) \wedge \bigwedge\{\neg(\varphi, \Delta \leq \psi) \mid \Delta \subseteq \mathcal{L}, \psi \in \Gamma\}$$

from Definition 3.4.3. We simply derive the rules of $PLFG$ in $NFLG$. For example, the first rule is already part of $NFLG$ and the second rule follows by the definition of strict ground in terms of weak ground by conjunction elimination. The other rules in the first row are equally straightforward. The only more complex derivation is the one of the long rule in the last row, called *reverse subsumption* by Fine. We don't give the proof here, as it spans two pages. But note that the side premises together allow us to derive the second conjunct $\bigwedge\{\neg(\varphi, \Delta \leq \psi) \mid \Delta \subseteq \mathcal{L}, \psi \in \Gamma\}$ of the definiens from Definition 3.4.3. □

3.7 Conclusion

Let's take stock. In this chapter we've developed a semantics for (iterated) ground based on the view that ground ground ground and we've given a sound logic for this semantics. In this conclusion, we wish to point out a few ways in which we might reasonably extend the semantics and logic that we've developed in this chapter.

First, obviously it would be nice to have a completeness proof for the logic and indeed this is feasible though tedious. Let's sketch how this result can be achieved. The idea would be to prove completeness by means of a canonical model construction. The idea would be that for a given set $\Gamma \subseteq \mathcal{L}$ of some language of ground, we first construct a state space $\mathcal{S}_\Gamma$, where

- $D_\Diamond = \wp(\mathcal{C})$ is the set of all sets of individual constants of $\mathcal{L}$,
- $D_@ = \{\{c_1, c_2 \mid \Gamma \vdash_{NFLG} c_1 = c_2\} \mid c_1, c_2 \in \mathcal{C}\}$ is the set of all equivalence classes terms that are provably identical according to Γ,
- $S_\Diamond = \{\Delta \subseteq \mathcal{L} \mid \text{for all } \varphi \in \Delta,\ \varphi \text{ is a literal of } \mathcal{L}\}$ is the set of all sets of *literals* of $\mathcal{L}$, where a literal is an atomic formula or its negation,
- $D_@ = \{\Delta \subseteq \mathcal{L} \mid \Delta \in D_\Diamond\ \Gamma \vdash_{NFLG} \varphi \text{ for all } \varphi \in \Delta\}$ is the set of all sets of literals that are provably true according to Γ,
- and finally $\prod = \bigcup$.

It is easily checked that this is indeed a state space according to Definition 3.5.1.

Next, we would define a *canonical intepretation* $\mathcal{I}_\Gamma$ for Γ, where

- $\delta(c) = \{c' \in \mathcal{C} \mid \Gamma \vdash c = c'\}$;
- $v^+(P)(\delta(c_1), \ldots, \delta(c_n)) = \{\Delta \mid \Gamma \vdash_{NFLG} \Delta \leq P(c_1, \ldots, c_n)\}$; and
- $v^-(P)(\delta(c_1), \ldots, \delta(c_n)) = \{\Delta \mid \Gamma \vdash_{NFLG} \Delta \leq \neg P(c_1, \ldots, c_n)\}$.

It can then be, somewhat more tediously, be checked that this is indeed a negatively adequate interpretation of $\mathcal{L}$ over $\mathcal{S}_\Gamma$ according to Definition 3.5.4. Then we'd have to show that a formula is true in this state space under this interpretation iff it is provable from Γ:

$$(\mathcal{S}_\Gamma, \mathcal{I}_\Gamma) \vDash \varphi \text{ iff } \Gamma \vdash \varphi$$

From this the claim would follow by standard arguments. But this is where things get hairy. The basic proof idea would be that we first show that every true sentence φ can be brought into a provably factually equivalent disjunctive normal form, where a sentence is in disjunctive normal form iff it is of the form

$$\bigvee\{(\bigwedge \Delta_i \mid \Delta_i \subseteq \mathcal{L} \text{ for } i \in I \text{ such that for all } \varphi \in \Delta \text{ are literals}\}.$$

More formally, we'd want to prove that for all $\varphi \in \mathcal{L}$ there is a $\theta \in \mathcal{L}$ in disjunctive normal form, such that $\Gamma \vdash \varphi \fallingdotseq \psi$. This is indeed feasible. By Lemma 3.5.18, all the steps that are usually used to prove the disjunctive normal form theorem are factual equivalence preserving. Moreover, by the properties (i.a-b,ii,) and (iii) of Lemma 3.5.17, we can successively eliminate occurrences of the ground operator $\leq$ in favor the conjunctions of its grounds or their negations, while preserving factual equivalence. In other words, we simply use the provable factual equivalence $(\Delta \leq \varphi) \fallingdotseq \bigwedge \Delta$ to get reduce sentences with the ground operator to sentences without, while preserving factual equivalence.

Then we would have to show that provability from Γ coincides with truth in $\mathcal{S}_\Gamma$ under $\mathcal{I}_\Gamma$ for all sentences in disjunctive normal form and infer the desired claim by the (provable) fact that (provable) factual equivalence means sameness of exact verifiers (in all models that satisfy Γ).[16] But here the problems start. Since our logic is infinitary, standard infinitary issues arise. Essentially, the problem is that we have to keep a tight control on the size of our languages of ground, in the sense that we need to limit the number of individual constants and the size of sets that we form infinite conjunctions, disjunctions, and statements of ground from.[17]

One kind of result that can indeed be established using the methods described by Green [56] and the strategy described before (although we will not do this here) is that if $\mathcal{L}$ is a language of ground with countably infinitely many individual constants and where all infinite conjunctions and disjunctions as well as all statements of ground are formed from countably infinite sets, then $NFLG$ is complete with respect to our semantics for countable premise sets: For all countable $\Gamma \subseteq \mathcal{L}$ and all $\varphi \in \mathcal{L}$, $\Gamma \vdash_{NFLG} \varphi$ iff $\Gamma \vDash \varphi$. We do not know how far this result can be extended, but we believe that some more fruitful work can be carried out in this direction, which will eventually shed more light on the infinitary nature of ground.

It would also be nice to extend our framework to allow for quantification. But here the infinitary issues intensify: As we've mentioned before, quantification in the context of ground arguably requires us to introduce multiary predicates in the sense of Oliver and Smiley [105]. Moreover, we'd have to keep an even tighter control over the size of our expressions. For example, we'd have to make sure that we only allow for quantifier blocks of finite length, if we harbor any hope for completeness. In any case, quantifiers in the context of ground are an exciting topic and we suspect that a lot of interesting research can be carried out in this direction as well.

Finally, a natural thought would be to *modalize* the semantics in order to allow for us to interpret statements about necessity *de re* in the framework. In particular, it would be desirable to validate the commonly accepted inference

$$\frac{\Gamma \leq \varphi}{\Box(\bigwedge \Gamma \to \varphi)}$$

[16]To prove completeness by means of disjunctive normal form is the proof idea that Fine uses to establish completeness for his pure logic of ground [44] and the logic of analytic entailment [37] over comparable semantic.

[17]For an overview of the problems that arise for completeness proofs in infinitary propositional languages, see [69, p. 30–54].

in our framework. An approach that immediately suggests itself would be to distinguish in the definition of a state space not only between the possible states $S_\Diamond$ and actual states $S_@$, but to have a set of states S_w for every possible world $w \in \mathcal{W}$, where w is the set of possible worlds. Then we'd just have to make sure that our interpretations respect classical (or negative free) logic at every possible world and adjust our definitions of exact verifiers and falsifiers accordingly, and we could then define truth of a non-modal sentence at a world $w \in \mathcal{W}$ to be that there is an exact verifier of the sentence that is also a member of S_w. Then could define truth of modal sentences of the form $\Box\varphi$ simply to be truth at every world. And on this approach we could likely indeed validate the above inference. Remember that a necessary condition for weak ground under our above semantics was to have that all *possible* states that are (fusions of) the exact verifiers of the grounds are exact verifiers of the groundee. And this property will not change across worlds, and so if this is indeed the case, then as soon as the grounds will be true at a world, so will be the groundee. But also this approach has limits. Note that again we just said *when* a modal statement is true in the kind of structure that we've just sketched, and not *why* it is true. In other words, we would like to have the exact verifiers and falsifiers of modal statements to properly account for modality in the present framework. But giving the exact verifiers and falsifiers of modal statements is something that is, at least at the present stage of research, out of reach. Again, we conjecture that a lot of very interesting research will be carried out in this direction.

Finally, we'd like to sum up what we've achieved toward the goal of this dissertation. In this chapter, we've presented an intuitive semantic framework in which we can interpret (iterated) ground as transmission of exact truthmakers in the weak sense of ground and irreversible transmission of exact truthmakers in the strict sense of ground. Moreover, by implementing negative free logic, we have achieved this in a way that is very much in the spirit of **Contingentism**: our framework has no existential commitments, since we allow for some objects in the semantics to not actually exist. Thus, if we were to extend the framework to account for modal statements as well as statements of ground, we are in the perfect position to implement **Contingentism** *full on*: we can allow for it to be possible that some objects possibly don't exist. But this is subject for future research. In the next chapter, we will develop a property theory in this framework and in the conclusion we will bring the framework to bear on the question of explicating **EGH**. But for now, let us point out that the semantic framework that we've discussed in this chapter is both formally precise and philosophically fruitful. In this conclusion, we've already pointed out a range of questions that can likely be fruitfully addressed in the present framework. And in the preface we already pointed out that Fine has brought this framework to bear on questions ranging from counterfactuals to intuitionistic logic, not to mention the origin of the semantics in the context of relevance logic. For more on the possible applications of this semantic framework, see [36]. The point here is that the semantic framework that we've brought to bear in this chapter is arguably (philosophical) fruitful in Carnap's sense.

Chapter 4

How To Distinguish Necessarily Equivalent But Distinct Properties

4.1 Preface

In this chapter, we will develop a hyperintensional theory of properties in the semantic framework that we used in the previous chapter. More specifically, we'll use the same idea that we've used to interpret predicates in this framework to develop a theory of properties. Intuitively, an n-ary predicate symbol expresses a relation. On the semantics of the last chapter, we've interpreted predicate symbols by means of functions that assign to every n-tuple of objects a set of possible states of affairs. Intuitively, this set corresponds to a disjunctive list of exact conditions for the objects to stand in the relation expressed by the predicate symbol: a sequence of n-objects stands in the relation expressed by the property symbol under an interpretation iff at least one of these possible states actually obtains. In this chapter, we will *reify* this semantics: we will simply understand properties as arbitrary such functions. We argue that this gives us a natural and philosophically adequate theory of fine-grained properties, which is perfectly suited to interpret essential properties as properties grounded in the identity of things.

4.2 Introduction

This chapter is about necessarily equivalent, but intuitively distinct properties. We call two properties Φ and Ψ *necessarily equivalent* iff for all objects x it is (metaphysically) necessary *de re* that x exemplifies Φ iff x exemplifies

Ψ. The history of philosophy is full of examples of necessarily equivalent but intuitively distinct properties. To name just a few examples, consider the properties of:

- being (spatially) extended and being colored [12],
- being triangular and being trilateral [54],
- being a round square and being a married bachelor [95], and
- being self-identical and being a member of one's singleton [39].[1]

In all of the above cases, one may plausibly hold that for every object it's necessary *de re* that it exemplifies the one property iff it exemplifies the other. For example, we might hold that for every object it's necessary *de re* that the object is triangular iff it is trilateral, because it is a theorem of Euclidean geometry that every polygon with three angles has three edges and vice versa. Or for another example, we might hold that for every object it is *vacuously* necessary *de re* that the object is a round square iff it is a married bachelor, because for every object it's necessary *de re* that it is not a round square and it's necessary *de re* that it is not a married bachelor. Similar arguments may be given for the other property pairs. But intuitively, in all of these cases the two properties are distinct. For example, being extended means to occupy a specific region of space, while being colored means to have some color. And certainly, these two things are different. For another example, being triangular means to have three angles, while being trilateral means to have three sides. And again, these two things are intuitively different. Analogous arguments may be given for the other property pairs. The aim of this chapter is to develop a property theory that can account for these intuitions.

A theory of properties that can distinguish between necessarily equivalent properties is called a *hyperintensional theory of properties*. Remember that we call a context *hyperintensional* iff in the context the substitution of necessary equivalents need not preserve truth-value [29, 103]. Thus, according to hyperintensional theories of properties, talk of properties creates hyperintensional contexts. To illustrate, on the relevant hyperintensional conception of properties, we might plausibly hold, for example, that

- *being self-identical is an essential property of Socrates*

is true, while

- *being a member of his singleton is an essential property of Socrates*

[1] As in our introduction, we take gerundives, like "being extended" and "being colored," as *canonical property designators*: expressions we normally use to denote to properties [125].

is false [39]. Or, for another example, we might plausibly hold that

- *being a round square is a geometric property*

is true, while

- *being a married bachelor is a geometric property*

is false. Thus, especially when we speak of higher-level properties of properties, such as being an essential property or being a geometric property, we have good intuitive reasons for developing a hyperintensional theory of properties.

More generally, hyperintensional property theories have applications all over metaphysics and the philosophy of language, since higher-order properties of properties often play an important role in these disciplines. For example, hyperintensional theories of properties can help us understand:

- the distinction between *essential* and *accidental* properties [39],
- the distinction between *fundamental* and *non-fundamental* properties [123],
- the distinction between *intrinsic* and *extrinsic* properties [3], or
- properties of *impossible* and *fictional entities* [113].

Thus, developing a hyperintensional theory of properties is not an isolated endeavor, but rather a well-connected project with useful applications in different areas of philosophy.

Given these applications, we can formulate some natural desiderata for a hyperintensional theory of properties. Intuitively, a hyperintensional theory of properties should satisfy the following conditions:

Grainedness. The theory should give us an adequate identity criterion for properties, which individuates properties:

(a) neither too coarsely,

(b) nor too finely.

Ontology. The theory should tell us what kind of objects properties are.

Applications. The theory should account for its applications, especially the applications mentioned above.

Let's discuss these conditions a bit further. The first condition **Grainedness** contains two components. First, it demands that a hyperintensional theory of properties give us an *identity criterion* for properties. An identity criterion for properties is a law of the following form [136]:

- For all properties Φ and Ψ, $\Phi = \Psi$ iff $C(\Phi, \Psi)$.

Here, $C(\Phi, \Psi)$ is a *condition* on the two properties Φ and Ψ.[2] Thus, in other words, a hyperintensional theory of properties should give us a necessary and sufficient condition for two properties to be identical.[3] This demand can be motivated by Quine's dictum "no entity without identity" [106, p. 23].

The second condition in **Grainedness** essentially says that a hyperintensional theory of properties had better really be a proper *hyperintensional* theory of properties. To illustrate how this condition might be failed to be satisfied, let's consider some counterexamples. According to an *extensional property theory*, two properties are identical iff they are materially equivalent. Thus, such a theory would give us the following criterion of identity for properties:

- For all properties Φ and Ψ, $\Phi = \Psi$ iff for all objects x (x exemplifies Φ iff x exemplifies Ψ).

But such a theory would already identify any pair of coincidentally coextensive properties, such as being a chordate and being a renate, to take Quine's famous example [115]. This will certainly not do for a hyperintensional theory of properties, as this would identify even more properties than the problematic property pairs listed above. Extensional theories of properties are what we may call *ultra-coarse*. Next, according to an *intensional property theory*, such as the theory **IPT** from our introduction, two properties are identical iff they are necessarily coextensive. Thus, on such a theory we would get the following identity criterion for properties:

- For all properties Φ and Ψ, $\Phi = \Psi$ iff $\Box$for all objects x(x exemplifies Φ iff x exemplifies Ψ).

Here the symbol '$\Box$' standardly expresses metaphysical necessity. Thus, on an intensional theory of properties we can distinguish between being a chordate and being a renate, since it is certainly metaphysically possible for an object to be a chordate without being a renate or vice versa. But also on such a theory, properties are individuated to coarsely. In particular, all the property pairs listed above are identified on such a theory: we get, for example, that being a married bachelor just *is* being a round square. And following our motivating intuitions also this will not do. Intensional property are not as coarse as extensional property theories, but still too coarse for the

[2]There is a discussion on what conditions we might want to put on identity criteria and in particular the condition C [61, 80, 136]. Here we follow the standard intuitive view that a criterion of identity should be *metaphysically necessary* and thus true, and that C should at least express an *equivalence relation*, i.e. a reflexive, symmetric, and transitive relation on the properties. But see the arguments in [80].

[3]In this chapter, we restrict ourselves to what Williamson [136, p. 144–48] calls *one level* criteria of identity for properties. What Williamson [136, p. 144–48] calls *two level* criteria of identity, or more traditionally *abstraction principles*, could also be applied to properties in interesting ways, but we shall not do so in this chapter.

applications we have in mind. They are what we might call *medium-coarse*.

On the other side of the spectrum, according to a *syntactic property theory*, properties are just (interpreted) predicates. For example, we would identify the property of being blue with the predicate "x is blue" under its intended interpretation. Thus, on such a theory two properties are identical iff they syntactically identical:

- For all properties Φ and Ψ, $\Phi = \Psi$ iff Φ and Ψ are identical as expressions.

But such a theory of properties is *too* fine-grained. For example, we might want to identify the property of being a bachelor and being an unmarried man. But on a syntactic theory of properties the two would be distinct, since "x is a bachelor" and "x is a man" are distinct syntactic entities. Thus, we might call syntactic property theories *ultra-fine*.

We're looking, however, for a theory of properties that lies somewhere between medium-coarse and ultra-fine. In other words, we're looking for a theory that is *medium-fine*. If we may illustrate this with a diagram, this is the idea:

$$\underbrace{\text{Extensional} < \text{Intensional}}_{\text{too coarse}} < \underbrace{\text{Hyperintensional}}_{\textit{just right!}} < \underbrace{\text{Syntactic}}_{\text{too fine}}$$

We're looking for a theory of properties that falls just in the right place with regards to grain. This might seem like an almost trivial demand on theories of properties, but as we shall argue, most hyperintensional theories of properties on the market actually violate this condition in one direction or the other, and sometimes even in both directions.

The second condition **Ontology** requires us to say what *kind* of object properties are. Syntactic theories of properties already give us an example for what an answer might look like: according to syntactic theories, properties are *predicates*. For another example, on the standard intensional theory of properties **IPT**, properties are *intensions*, i.e. functions from possible worlds to sets of objects. What motivates this condition is that we wish to implement our theory of properties in our overall metaphysics. Moreover, the condition ties in with the first component of the first condition **Grainedness**: if we say what kinds of objects properties are then we presumably can derive from this a natural identity criterion for properties. This is again illustrated by syntactic theories of properties, where we can infer that since properties are predicates they should be individuated as such, giving us the above identity criterion for properties. Similarly, on the standard intensional theory of properties **IPT**, properties are functions and thus individuated as such. Remember that functions are identical iff they agree on all their arguments, thus we naturally get the following identity criterion for properties according to **IPT**:

- For all properties Φ and Ψ, $\Phi = \Psi$ iff for all possible worlds $w \in \mathcal{W}$ $\Phi(w) = \Psi(w)$.

Moreover, the condition **Ontology** ties in with Carnap's idea of an explication [17, p. 1–18]: by saying what kind of object properties are, we're effectively replacing the intuitive concept of a property with a more precise concept in a previously understood ontology.

The third condition **Applications** is almost self explanatory: as we've pointed out above, the main motivation for developing a hyperintensional theory of properties are its philosophical applications; and if a candidate hyperintensional theory of properties fails to account for these applications, then the whole motivation for developing the theory vanishes—the theory becomes, if we may put it so harshly, *pointless*.

We can sum up the point of these conditions in terms of Carnap's concept of an *explication* of a concept. According to Carnap: "the task of *explication* consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second" [17, p. 3, emphasis in the original]. In a given explication, Carnap calls the explicated concept the *explicandum*, and the explicating concept the *explicatum*. Before giving an explication of a concept, so Carnap, we should try to make the explicandum as precise as possible: for example, we should give an informal definition of the explicandum, we should give examples of the correct application of the concept, or we should discuss its logic [17, p. 4–5]. Only then, we should attempt to find a proper explicatum. Once we've determined an explicatum, according to Carnap, the quality of the explication can be judged by four criteria: (1) the *similarity* between the explicatum and the explicandum; (2) the *exactness* of the explicatum; (3) the *fruitfulness* of the explicatum; and (4) the *simplicity* of the explicatum [17, p. 5–8].

The idea is that we understand the project of giving a hyperintensional theory of properties as a Carnapian problem of explication for a hyperintensional concept of properties that agrees with our intuitions in the cases of the property pairs mentioned at the outset. In this introduction, we've given different intuitive examples that determine the correct use of this hyperintensional concept of properties. And the three conditions **Grainedness**, **Ontology**, and **Applications** correspond to Carnap's conditions (1–3) when applied to the concept of properties in question.

4.2.1 Overview of the Chapter

This chapter has two parts, one negative and one positive. In §4.3, we will discuss the main proposals for hyperintensional property theories on the market: the *impossible worlds theory* and the *structured properties theory.* We will argue that neither of the two theories fares well with regard to all three of our desiderata for hyperintensional property theories. Then, in §4.4, we will develop our own hyperintensional property theory and argue that this theory does not fall prey to the problems of the impossible worlds theory or the structured properties theory. Indeed, we shall argue that our theory overcomes the problems of these theories in a natural and fruitful way.

4.3 Hyperintensional Property Theories

A wide range of property theories can be found in the literature, but not always are properties taken to be the subject in their own right.[4] Therefore, we will, for the most part, *extract* theories of properties from approaches that may not be primarily intended to provide a theory of properties. The two main contenders for a hyperintensional theory of properties that we come up with this way are the *impossible worlds theory* and the *structured properties theory.* Let's go through these theories in turn.

4.3.1 The Impossible Worlds Theory

The impossible worlds theory uses *impossible* worlds in addition to possible worlds to distinguish between necessarily equivalent but intuitively distinct properties. Thus the framework in which the theory is formulated is an extension of the usual possible worlds framework. Intuitively, impossible worlds correspond to the ways the world could *not* have been, just like possible worlds correspond to the ways the world *could* have been.[5] Given the concept of possible worlds, it is plausible to postulate the following principle of *plenitude*:

Plenitude for Possible Worlds. For every way the world could have been there is a possible world that corresponds to this way.

Such a principle is, for example, endorsed by Lewis [83] and others. Analogously, given the concept of impossible worlds, it is plausible to postulate a

[4]Notable exceptions are [8, 9, 98], [83, p. 50–69], and, of course, [140]. For a general discussion of philosophical property theories, see [130].

[5]For a more comprehensive discussion of the concept of possible worlds, see Menzel [97], and for a more comprehensive discussion of the concept of impossible worlds, see [14].

corresponding principle of plenitude for impossible worlds:

Plenitude for Impossible Worlds. For every way the world could *not* have been there is an *im*possible world that corresponds to this way.

Such a principle is, for example, endorsed by Priest [112, 113] and others.

And just like possible worlds are inhabited by possible objects or *possibilia*, impossible worlds are inhabited by impossible objects or *impossibilia*. The properties of possible objects are constrained by what is possible: given **Plenitude for Possible Worlds** and the concept of possible worlds, it is plausible to say that possibilia can only have properties that they can possibly exemplify. Thus, for example, there is no possible object in any possible world that is a round square or a married bachelor. Impossibilia, in contrast, are not constrained in this way: given **Plenitude for Impossible Worlds** and the concept of impossible worlds, it is plausible to say that impossibilia can have all the properties we like—even inconsistent ones. So there is an impossible object in some impossible world that is a round square and there is an an impossible object in some impossible world that is a married bachelor. Of course, impossible worlds may also be inhabited by some possible objects, but possible worlds are never inhabited by impossible objects.

In the following, we denote the set of possible worlds by $\mathcal{W}$ and the set of impossible worlds by $\mathcal{I}$.[6] We denote the set possible or impossible objects that exist at a possible or impossible world $w \in \mathcal{W} \cup \mathcal{I}$ by $\mathcal{D}_w$.

Remember that an *intension* is a function i that assigns to every world $w \in \mathcal{W}$ an *extension* $i(w) \subseteq \mathcal{D}_w$, i.e. a set of possible objects that inhabit the world. On the intensional property theory, properties are identified with intensions and the extension that is assigned by a property to a world intuitively corresponds to the set of things that exemplify the property at that world. The impossible worlds theory of properties generalizes this idea. Let's take an *hyperintension* h be a function that assigns to every world $w \in \mathcal{W} \cup \mathcal{I}$, possible or impossible, a set $h(w) \subseteq \mathcal{D}_w$ of possible or impossible objects that inhabit the world. Correspondingly may call such a set an *impossible hyperextension*. The idea of the impossible worlds theory of properties is that, just like the intensional theory of properties **IPT** takes properties to be intensions, we may alternatively understand properties as *pairs* of hyperintensions:

[6] Priest [113] distinguishes more finely between *open* and *closed* worlds. The difference is that open worlds need not decide every fact of the matter, while closed worlds do. In other words, open worlds are incomplete, while closed worlds are always complete. For Priest only closed worlds are impossible worlds and open worlds have their own ontological category. But since the actual world could intuitively not have been incomplete, we also count open worlds as impossible worlds.

Impossible Worlds Theory (IWT). For all Φ, Φ is a property iff Φ is a a pair of functions (Φ^+, Φ^-) such that for every possible or impossible world $w \in \mathcal{W} \cup \mathcal{I}$:

(i) Φ^+ assigns an extension $\Phi^+(w) \subseteq \mathcal{D}_w$ to w, and

(ii) Φ^- assigns an anti-extension $\Phi^+(w) \subseteq \mathcal{D}_w$ to w,

such that for all $w \in \mathcal{W}$, both

(a) $\Phi^+(w) \cup \Phi^-(w) = \mathcal{D}_w$, and

(b) $\Phi^+(w) \cap \Phi^-(w) = \emptyset$.

For example, on this theory we could identify the property of being red with the function pair (red$^+$,red$^-$) such that red$^+$ assigns to every world, possible or impossible, the set of red things in the world and red$^-$ assigns to every world the set of non-red things at the world. Such a theory is, more or less explicitly, endorsed by Priest [113], Ripley [120], and Jago [64].

The intuition behind **IWT** is that, analogously to the intensional property theory **IPT**, an object exemplifies a property at a world, possible or impossible, iff it is in an element of the extension of the property at the world. But, in contrast to the intensional property theory, on the impossible worlds theory, a property also gives us for every world a set of objects that intuitively *fail* to exemplify the property. This set is called the *anti-extension* of the property. On the intensional property theory **IPT**, in contrast, the anti-extension of a property Φ at a world $w \in \mathcal{W}$ is usually simply defined as $\mathcal{D}_w \setminus \Phi(w)$, i.e. as the relative complement of the extension of the extension $\Phi(W)$ of the property at he world with respect to the things that inhabit the world. Consequently, it is impossible for a possible object $x \in \mathcal{D}_w$ that inhabits some possible world $w \in \mathcal{W}$ both to exemplify and fail to exemplify some property Φ at the world. Nor is it possible for some possible object $x \in \mathcal{D}_w$ in some world $w \in \mathcal{W}$ neither to exemplify Φ nor fail to exemplify Φ at w. By conditions (a) and (b) of **IWT**, the impossible worlds theory agrees with this: at every *possible* world $w \in \mathcal{W}$ we get for every property $\Phi = (\Phi^+, \Phi-)$ that $\Phi^-(W) = \mathcal{D}_w \setminus \Phi^+(w)$. Thus, also according to **IWT**, it is impossible for an object at some *possible* world both to exemplify and fail to exemplify some property at that world. And similarly, it is impossible for some object at some *possible* world neither to exemplify not fail to exemplify some property at that world. But at an *im*possible world, anything goes. It is possible for some property $\Phi = (\Phi^+, \Phi^-)$ that at some impossible world $w \in \mathcal{I}$ some impossible object $x \in \mathcal{D}_w$ is such that both $x \in \Phi^+(w)$ and $\Phi^-(w)$ or that x is such that neither $x \in \Phi^+(w)$ nor $x \in \Phi^-(w)$. For example, at some impossible world $w \in \mathcal{I}$ there may be some ball $b \in \mathcal{D}_w$ that is both red and non-red at that world, and thus $b \in \text{red}^+(w)$ and $b \in \text{red}^-(w)$. Or there may be an impossible world $w \in \mathcal{I}$ with some ball $b \in \mathcal{D}_w$ that is neither red nor not-red at that world, and thus $b \notin \text{red}^+(w)$ and $b \notin \text{red}^-(w)$

To avoid confusion between failing to exemplify a property at a world in the sense of being a member of the anti-extension of the property at the world, and failing to exemplify a property at a world in the sense of not being a member of the extension of the property at the world, since the two notions come apart in the framework of impossible worlds, we call the former relation *anti-exemplification* and reserve the natural language expression failing to exemplify a property at a world for the latter relation. To sum up, on the impossible worlds theory **IWT**, exemplification and anti-exemplification are analyzed as follows:

Exemplification and Anti-Exemplification on IWT. For all properties Φ, for all worlds $w \in \mathcal{W} \cup \mathcal{I}$, and for all objects $x \in \mathcal{D}_w$:

- x *exemplifies* Φ at w iff $x \in \Phi^+(w)$, and
- x *anti-exemplifies* Φ at w iff $x \in \Phi^-(w)$.

And even though it is not possible for some possible world $w \in \mathcal{W}$ that some object $x \in \mathcal{D}_w$ both to exemplify and anti-exemplify a property, it is very well possible that for some impossible world $w \in \mathcal{I}$ that some object $x \in \mathcal{D}_w$ both exemplifies and anti-exemplifies some property. And similarly, even though it is not possible for some possible world $w \in \mathcal{W}$ that some object $x \in \mathcal{D}_w$ neither to exemplify not to anti-exemplify a property, it is very well possible that for some impossible world $w \in \mathcal{I}$ that some object $x \in \mathcal{D}_w$ neither exemplifies nor anti-exemplifies some property.

On the impossible worlds theory **IWT**, properties are pairs of functions and thus they are individuated as such. In particular, we get the following identity criterion for properties:

Property Identity on IWT. For all properties $\Phi = (\Phi^+, \Phi^-)$ and $\Psi = (\Psi^+, \Psi^-)$, $\Phi = \Psi$ iff for all worlds $w \in \mathcal{W} \cup \mathcal{I}$ we have $\Phi^+(w) = \Psi^+(w)$ and $\Phi^-(w) = \Psi^-(w)$.

As a consequence, **IWT** is indeed able to distinguish between necessarily equivalent properties. Take being colored and being extended as an example. As we've said, we can plausibly hold that for every object x it's necessary *de re* that (x exemplifies being colored iff x exemplifies being extended). And by the usual understanding of necessity *de re* as what's the case in every possible world, this means that in every possible world everything that is colored in the world is extended and vice versa. But in the framework of impossible worlds, there may very well be an *im*possible world where something is extended but fails to be colored or colored but fails to be extended. Indeed, by **Plenitude for Impossible Worlds**, there will be such a world, since the actual world can't be such that something is extended but not colored or colored but not extended. Hence, if we take the property of being colored in the natural way to be the pair (colored$^+$, colored$^-$) such that colored$^+$ assigns to every world, possible or impossible, the things that are colored at the world and colored$^-$ assigns to every world the set of things that are colorless at the world, and we take the property of being extended in the natural way to be the pair (extended$^+$, extended$^-$) such that extended$^+$ assigns to every world the things that are extended at the world and colored$^-$ assigns to every world the set of things that are unextended at the world, then we will get that the properties of being colored and being extended are distinct. For even though the function pairs (colored$^+$, colored$^-$) and (extended$^+$, extended$^-$) agree on all possible worlds, they go apart on some impossible worlds. By **Plenitude for Impossible Worlds** there is a world $w \in \mathcal{I}$ such that there is an $x \in \mathcal{D}_w$ such that $x \in$ colored$^+$ but $x \notin$ extended$^+$. Hence colored$^+ \neq$ extended$^+$ and thus (colored$^+$, colored$^-$) is distinct from (extended$^+$, extended$^-$). In other words, according to **IWT**, the properties of being colored and being extended are distinct—as they should be. We can distinguish between necessarily equivalent but intuitively distinct properties.

But there is a problem with this story. It seems that according to **IWT**, we can distinguish *any* two properties—and this is counterintuitive. Take, for example, the properties of being a bachelor and being an unmarried man. Intuitively, being a bachelor *means* being an unmarried man, as is supported by the commonplace claim that "all bachelors are unmarried men" and "all unmarried men are bachelors" are *analytic truths*: truths in virtue of the meaning of "bachelor" and "unmarried man." But according to the principle **Plenitude for Impossible Worlds** there will be an impossible world where some bachelor fails to be an unmarried man. Thus if we understand the property of being a bachelor in the natural way as the function pair (bachelor$^+$, bachelor$^-$), where bachelor$^+$ assigns to every world, possible or impossible, the objects that are bachelors at the world and bachelor$^-$ assigns to every world the set of objects that are no bachelor at the world, and if we take the property of being an unmarried man in the natural way to be the pair (unmarried man$^+$, unmarried man$^-$) such that unmarried man$^+$ assigns to every world the things that are unmarried men at the world and unmarried man$^-$ assigns to every world the set of things that are not unmarried man at the world, then we get that the properties of being a bachelor and being an unmarried man are distinct. For by the principle **Plenitude for Impossible Worlds** there will be an impossible world $w \in \mathcal{I}$ such that there is an $x \in \mathcal{D}_x$ is such that $x \in$ bachelor$^+$ and $x \notin$ unmarried man$^+$. Hence the two properties are distinct. Note that the argument is completely analogous to how we showed that according to **IWT** being extended and being colored are distinct. It seems that an analogous argument can be carried out for *any* property pair. This is certainly counterintuitive. For example, the properties of being a bachelor and of being an unmarried man, or the properties of being a ball and of being a ball or a ball, are intuitively identical—yet they are arguably distinguished by the impossible worlds theory **ITW**.

The obvious candidate for the culprit here is the principle **Plenitude for Impossible Worlds**. In the arguments above, we've used the principle to establish that for every property, understood as a pair of hyperintensions, there are impossible worlds where these hyperintensions go apart and thus the properties have to be different. But to fix this problem, we have abandon the principle **Plenitude for Impossible Worlds** in its full generality: we have to somehow restrict the principle. Given the intuitive concept of an impossible world, however, this seems difficult. Why should something be "too impossible" for there to be an impossible world that corresponds to this impossibility. For example, why shouldn't there be an impossible world where some bachelor is married? It seems that there is no non-*ad hoc* answer to this. We take it that this is a serious objection to the impossible worlds theory and we don't see how it can be overcome.[7]

To sum up, the impossible worlds theory **ITW** does give us an identity criterion for properties but the identity criterion is too fine grained: the theory violates our condition (b) of our desideratum **Grainedness**. The theory does tell us what kind of objects properties are, they are pairs of hyperintensions, and thus the theory fares quite well with regard to our third desideratum **Ontology**. But also the status of the theory with regard to **Applications** is dubious. It seems unclear, for example, how the hyperintensional concept of essential properties that we're interested in in this dissertation should be analyzed in the framework of impossible worlds. Clearly, the modal analysis **PWA** will not work. For even though **IWT** will allow us to distinguish between the two properties of being self-identical and being a member of one's singleton, for example, it will still be the case for every object that the object is self-identical at a possible world iff it is a member of its singleton there. And hence, it will still be the case that both properties, though distinct, are essential properties of any object according to **PWA**. Moreover, given the principle **Plenitude for Impossible Worlds** it is unclear how **PWA** could be fixed in terms of impossible worlds. If we would say, for example, that a property Φ is essential to an object x iff x exemplifies Φ at every world, possible or impossible, where x exists, then we would get a trivial concept of being an essential property: given **Plenitude for Impossible Worlds**, we will be able to find a counterexample to any claim of a property being essential to an object. Thus, **IWT** does not fare particularly well with regard to **Applications** either. And for these combined reasons,

[7] We can of course construct models in a more formal sense that satisfy our intuitive desiderata outlined above, but that is beside the point. We can similarly restrict the principle **Plenitude for Impossible Worlds** in such a way that we validate our desiderata. But the question here is for a *justification* of this restrictions, which is not *ad hoc*. And we don't see how such a restriction can be given.

we will abandon the approach.

4.3.2 The Structured Properties Theory

The structured properties theory uses a concept of property structure to distinguish between necessarily equivalent but intuitively distinct properties. There are two versions of the structured properties theory in the literature: one *algebraic* and one *quasi-syntactic*. Let's discuss them in turn.

The algebraic variety of the structured properties theory is, for example, endorsed by Bealer [8], McMichael and Zalta [94], Bealer and Mönnich [9], and Menzel [98]. These authors propose to understand properties as the members of *property algebras*, which consist of a set of properties $\mathbb{P}$ that is closed under algebraic operations among properties, such as *property conjunction* **conj**, *property disjunction* **disj**, *property negation* **neg**, and so on. Here we shall not define these algebras explicitly, since our point is independent of the concrete properties of these algebras. Let's assume that the concept is defined in some way. The only point that is important is that these algebras are defined by means of *characteristic equations* for the members of $\mathbb{P}$ with respect to the algebraic operations on $\mathbb{P}$. For example, they might include a law of the form:

- For all $\Phi, \Psi \in \mathbb{P}$, $\mathbf{conj}(\Phi, \Psi) \neq \Phi$ and $\mathbf{conj}(\Phi, \Psi) \neq \Psi$.

A property algebra is then defined as any system of objects with the corresponding operations that satisfies these characteristic equations. And a property is then simply defined as a member of some property algebra:

Algebraic Structured Property Theory (ASPT). For all Φ, Φ is a property iff $\Phi \in \mathbf{P}$, where $\mathbb{P}$ is the underlying set of some property algebra.

Then, two properties are said to be identical iff they are identical as members of their property algebra. In some cases, this can be shown by means of the characteristic equations of the property algebra. For example, if we assume the above law, then we can show that for all properties Φ, then $\mathbf{conj}(\Phi, \Phi) \neq \Phi$. This would then indeed mean that **ASPT** is a hyperintensional theory of properties, since intuitively, the conjunctive property $\mathbf{conj}(\Phi, \Psi)$ is necessarily equivalent to its conjunct property Φ. We could, of course, also postulate other characteristic equations to suit our intuitive needs and make our theory exactly as fine-grained as we'd like. For example, we could postulate the following equations instead of the above one:

- For all $\Phi \in \mathbb{P}$, $\mathbf{conj}(\Phi, \Phi) = \Phi$.
- For all $\Phi, \Psi \in \mathbb{P}$, if $\Phi \neq \Psi$, then $\mathbf{conj}(\Phi, \Psi) \neq \Phi$ and $\mathbf{conj}(\Phi, \Psi) \neq \Psi$.

Then, we'd get that for all properties Φ, $\mathbf{conj}(\Phi, \Phi) = \Phi$. But this will only work for those properties that are described by the characteristic equations of the property algebra. Ultimately, property identity will be *primitive* in **ASPT**: two properties Φ and Ψ are identical iff they are identical objects. This means that **ASPT** does not fare well with regard to the first component of our desideratum **Grainedness**: the theory does not give us any meaningful identity criterion for properties.

To make things worse, **ASPT** does not fare well with regard to the criterion **Ontology** either. Note that a property algebra is defined as *any* system of objects with the corresponding operations that satisfies these characteristic equations, and a property is defined as an object in *some* property algebra. But this is counterintuitive: It might turn that some collection of numbers, or possible monkeys, or the like can be equipped with operations that satisfy the characteristic equations of a property algebra. But intuitively, properties are not numbers, possible monkeys, or the like. But more importantly, a consequence of this definition is that *anything* can be a property according to **ASPT** and thus the theory does certainly not tell us what kinds of objects properties are. This makes it difficult, if not impossible, to ascertain the metaphysical merit of **ASPT**. The point is that **ASPT** seems to rather describe the *structure* of properties, in the sense of the relations they bear to each other. But **ASPT** doesn't tell us what properties are—it doesn't give us any meaningful identity criterion nor does it tell us what ontological category properties belong to.[8] For this reason, we will abandon the algebraic structured properties theory **ASPT**.

The quasi-syntactic variety of the structured properties theory, in contrast, is endorsed, more or less explicitly, by Cresswell [29] and King [70], and Lewis [83, p. 50–69] for some applications. This theory is based on the observation that we typically express properties by means of *property expressions*. Standardly, we may take property expressions to be *(complex) unary predicates* of some (relatively) *informal* language, such as "x is a bachelor," "x is an unmarried man," or "x is self-identical," for example.[9] These property expressions come with an intended interpretation. For example, we assume that 'x is a bachelor" expresses the property of being a bachelor, "x is an unmarried man" expresses the property of being an unmarried man, and "x is self-identical" expresses the property of being self-identical. Thus, we usually talk about properties by means of property expressions under an

[8]We don't wish to exclude that properties might belong to a primitive ontological category. But if there is a primitive ontological category of properties, then we'd arguably need an identity criterion for the objects in that category and **ASPT** does not provide this.

[9]But they could be any kind of expression that we use to refer to properties, such as gerundives like "being a man" for example. Here we shall focus on unary predicates, however, since this is the standard way of developing the view.

intended interpretation.

To obtain a hyperintensional theory of properties from this observation, we first subject the intuitive property expressions to a syntactic analysis. In the standard way of doing this, we would simply let the syntactic analysis of our language do this job for us. For example, we would analyze "x is a bachelor" and "x is self-identical" as primitive property expressions, but we would analyze "x is an unmarried man" as the complex conjunctive predicate "x is a man & x is unmarried," where "x is a man" and "x is unmarried" are in turn primitive predicates. The result would be a definition of an expression φ being a well-formed property expression (read unary predicate) formed (in order) from the expressions $\sigma_1, \ldots, \sigma_n$, which that range over predicates and logical operators. For example, we would get that "x is unmarried" is a well-formed property expression formed only from itself, and "x is a man & x is unmarried" would be a well-formed property expression that is formed from "x is a man," "&, " "x is unmarried." In other words, we would subject our informal language to a syntactic analysis and in this way transform it into a (more or less) *formal* language. In the following, we'll write $\varphi(\sigma_1, \ldots, \sigma_n)$ to indicate that φ a well-formed property expression that is formed (in order) from the expressions $\sigma_1, \ldots, \sigma_n$, which range over predicates and logical operators.

Next, we define the notion of an *interpretation* for the expressions that we've previously syntactically analyzed. Standardly, we would take the interpretation of a predicate symbol to be an intension, i.e. a function from possible worlds to the sets of objects at that possible world.[10] The extension assigned to a predicate symbol, then are the objects that at the world exemplify the property expressed by the predicate under the interpretation. In the case of the logical operators, we'd take the semantic value to be the corresponding logical operations. This step is usually taken care of by carefully laying out the standard semantics for the kind of (formal) language we're considering. Next, we'd somehow identify the *intended* interpretation ν that assigns the intuitively correct intensions and operations to the predicates and operators of our language. For example, ν would be such that ν("x is the man") is the function that assigns to every possible world the set of all men at that world, and ν("&") would be the logical operation of conjunction on intensions.[11]

Then, finally, putting the two steps together, proponents of the quasi-syntactic structured property theory project the syntactic structure of the property expressions on their semantic values. The result are quasi-

[10]Again there is room for maneuver here. If we can give some alternative notion of an interpretation of property expressions, then also this notion would give rise to a (different) quasi-syntactic structured property theory.

[11]More specifically, this would be the function $\wedge$ such that for all worlds $w \in \mathcal{W}$ and all intensions i, j, $(i \wedge j)(w) = i(w) \cap j(w)$.

syntactically structured sequences of semantic values that reflect the syntactic structure of the property expressions intuitively expressing them. We end up with the following theory of properties:

Quasi-Syntactic Structured Properties Theory (QSPT). For all Φ, Φ is a property iff Φ is ordered-tuple $(\nu(\sigma_1), \ldots, \nu(\sigma_n))$ of the semantic values $\nu(\sigma_1), \ldots, \nu(\sigma_n)$ under the intended interpretation ν of the syntactic components of a well-formed property expressions $\varphi(\sigma_1, \ldots, \sigma_n)$.

Versions of this theory are, more or less explicitly, endorsed by Cresswell [29] and King [70], and Lewis [83, p. 50–69] for some applications.

Since according to **QSPT** properties are n-tuples, they are individuated as such. In particular, we get the following identity criterion for properties according to the theory:

Property Identity in QSPT. For all properties $\Phi = (\nu(\sigma_1), \ldots, \nu(\sigma_n))$ and $\Psi = (\nu(\tau_1), \ldots, \nu(\tau_m))$, for $n, m \in \mathbb{N}$:

- $\Phi = \Psi$ iff $n = m$ and $\nu(\sigma_i) = \nu(\tau_i)$, for all $1 \leq i \leq n$.

Thus, the quasi-syntactic structured properties theory **QSPT** improves over the algebraic structured properties theory **ASPT** in the sense that it gives us an identity criterion for properties and tells us what kind of objects properties are—they are tuples of semantic values of property expressions. Hence, **QSPT** fares quite well with regard to our desideratum **Ontology** and the first component of **Grainedness**.

But we wish to argue that there are also problems for the the quasi-syntactic structured properties theory **QSPT**. In particular, it appears that **QPST** violates *both* condition (a) and (b) of **Grainedness**: the theory individuates some properties to coarsely and some properties to finely. For the first point, consider the properties of being self-identical and being contained in one's singleton. We would typically express these properties by means of the predicates "x is self-identical" and "x is contained in $\{x\}$". But both of these predicates are atomic and thus we would get that the property of being self-identical according to **QSPT** is simply the one-tuple (ν("x is self-identical")) and the property of being self contained in one's singleton would be (ν("x is contained in $\{x\}$")), where ν gives the intended interpretation for these predicates. Moreover, the intended interpretation of "x is self-identical" is standardly the intension which maps every world to the set of objects that are self-identical in the world, and the intended interpretation of "x is self-identical" is standardly the intension which maps every world to the set of objects that are contained in their singleton in this world. But since we've agreed that being self-identical and being contained in one's singleton are necessarily equivalent, these intensions will be *identical*, i.e. ν("x is self-identical")$=\nu$("x is contained in $\{x\}$"). But then it follows by **Prop-**

erty Identity in QSPT that also the properties of being self-identical and of being contained in one's singleton are identical—contrary to what we want to say intuitively. The point here is that **QSPT** apparently cannot distinguish between necessarily equivalent properties which are expressed by *atomic* property expressions. And the only way of fixing this problem in the framework of **QSPT** would be to devise an intended interpretation that assigns different semantic values to necessarily equivalent but intuitively distinct atomic property expressions. But what would then be the point of **QSPT** in the first place?

For the second point, consider the properties of being a bachelor and being an unmarried man. Standardly, after syntactic analysis, we would express these properties by the predicates "x is a bachelor" and "x is a man & x is unmarried." Thus, according to **QSPT**, the property of being a bachelor would be the one-tuple (ν("x is a bachelor")) and the property of being an unmarried man would be the triple (ν("x is a man"), ν("&"), ν("x is unmarried")), where ν again assigns the intended interpretation to these expressions. But these two tuples *cannot* be identical, since the one is a one-tuple and the other is a triple. Yet, intuitively, we might want to identify the properties of being a bachelor and being an unmarried man, as we've pointed out before. More generally, according to **QSPT** a property that is expressed by a conjunctive property expression can *never* be identical to a property that is expressed by an atomic property expression. But this seems counterintuitive: identity of properties should not be determined by the complexity of the expressions expressing them.

To sum up, also **QSPT** does not fare very well with regard to our desiderata: in the case of properties expressed by atomic property expressions, the theory falls back into the problems of the intensional property theory **IPT** and in the case of properties expressed by syntactically complex property expressions the theory behaves too much like a syntactic theory of properties. For this reason, we shall also abandon structured property theories for the purpose of this chapter.

4.4 The Exemplification-Criteria Theory

In this section, we'll develop our new hyperintensional theory of properties. We'll develop this theory in an informal version of Fine's semantic framework of exact truthmaker semantics, which we've discussed at length in the previous chapter. But let's briefly recall the basics of the the framework. First, we assume that we're given a non-empty set $D_\Diamond$ of *possible objects*, which intuitively correspond to all the objects that could possible have existed. We furthermore assume that we're given a non-empty set $S_\Diamond$ of *possible states* *(of affairs)*. Intuitively these possible states are possible states of objects being a certain way. Just like in the case of possible and impossible worlds, it makes sense to postulate a principle of plenitude for possible states:

Plenitude for Possible States. If $x_1, x_2, \ldots$ are objects, then for every non-disjunctive way that $x_1, x_2, \ldots$ could have been like, there is a possible state of $x_1, x_2, \ldots$ being that way.

For example, if we have a ball that can be colored in different ways, then there will be possible states corresponding to the different ways that the ball could have been colored: there is a possible state of the ball being red, a possible state of the ball being green, a possible state of the ball being blue, and so on.

Any collection of states may intuitively be *(mereologically) fused*. Hence, we assume that we have an operation $\prod : \wp(S_\Diamond) \to S_\Diamond$ of *fact fusion* such that:

- for all states $s \in S_\Diamond, \prod\{s\} = s$ ('idempotence'); and
- for all indexed families $(X_i \subseteq S_\Diamond)_{i \in i}$ of states,

$$\prod\{\prod X_i \mid i \in I\} = \prod \bigcup \{X_i \mid i \in I\} \text{ ('commutativity').}$$

In other words, fact fusion respects the laws of mereology. Thus, intuitively, if there is a state of the ball being red and a state of the ball being round, then there is also the fusion state $\prod$\{the ball being red, the ball being round\}=the ball being red and round. Instead of $\prod$\{the ball being red, the ball being round\}, we'll usually just write:the ball being red$\circ$the ball being round.

Finally, we assume a non-empty set of individuals $D_@ \subseteq D_\Diamond$ of *actual objects*, which intuitively corresponds to the objects that actually exist. And we assume that we have a non-empty set $S_@ \subseteq S_\Diamond$ of *actual states*, which intuitively corresponds to the actually obtaining states. Thus, if there actually exists some green ball b, then the possible state of the ball b being green is a member of $S_@$. For intuitive reasons, we assume that the actual states are upwards and downwards closed with respect to state fusion: if an arbitrary state fusion $\prod X$, for $X \subseteq S_\Diamond$ actually obtains, then all the members of X actually obtain, and if all the members of $X \subseteq S_\Diamond$ actually obtain, then $\prod X$ actually obtains. More precisely:

- for all $X \subseteq S_@, \prod X \in S_@$, and
- for all $X \subseteq S_\Diamond$, if $\prod X \in S_@$, then $X \subseteq S_@$.

Thus, if there is some ball such that the states of the ball being red and the ball being round actually obtain, then the fusion state of the ball being red and round actually obtains. And similarly, if the fusion state of the ball being red and round actually obtains, then also its part states of the ball being red and the ball being round actually obtain.

As we've shown in the last chapter, this framework can be used to give truth conditions for (iterated) statements of ground and Fine has shown in different papers that the semantics can be brought to bear on issues ranging from counterfactuals to intuitionistic logic [42, 44, 44]. Now we will show that in this framework, we can also define a reasonable notion of a property. And as a consequence, this notion of a property can be used in all of the previously mentioned applications.

The idea is that we take a property to be a pair of functions which assign to every object a set of possible states that correspond to the precise conditions for the object to exemplify the property and a set of states which correspond the precise condition for the object to fail to exemplify the property. We'll call the first set the *exemplification criteria* and the second set the *anti-exemplification criteria* of the property. More precisely, we get the following definition of a property in our framework:

The Exemplification-Criteria Theory (ECT). A property Φ is a pair (Φ^+, Φ^-) of functions such that for every $x \in D_\Diamond$:

- Φ^+ assigns a set $\Phi^+(x) \subseteq S_\Diamond$ of *exemplification criteria* to x, and
- Φ^- assigns a set $\Phi^-(x) \subseteq S_\Diamond$ of *anti-exemplification criteria* to x,

such that for all $X \subseteq S_\Diamond$:

(i) if $X \subseteq \Phi^+$ non-empty, then $\prod X \in \Phi^+$.

(ii) if $X \subseteq \Phi^-$ non-empty, then $\prod X \in \Phi^-$.

The conditions (i) and (ii) essential amount to saying that exemplification and anti-exemplification criteria can be overdetermining whether or not the object exemplifies the property: all the fusions of exact conditions are themselves exact conditions.

The idea of **ECT** is that the exact exemplification and anti-exemplification criteria of a property for an object are disjunctive lists of all the possible ways in which the object can exemplify the property or fail to exemplify the property respectively. Thus, an object exemplifies a property iff some of the exact exemplification criteria of the property for the object actually obtain and an object fails to exemplifies a property iff some the exact anti-exemplification criteria of the property for the object actually obtain. More precisely, we get the following definition of exemplification and anti-exemplification on our theory:

Exemplification and Anti-Exemplification in to ECT. For all properties Φ and all objects $x \in D_\Diamond$ we say:

- x *exemplifies* Φ iff $\Phi^+(x) \cap S_@$; and
- x *anti-exemplifies* Φ iff there is an $s \in \Phi^-(x)$ such that $s \in S_@$.

Thus, intuitively the property of being red will assign to a ball the set of only the state of the ball being red as its only exact instantiation criterion and it will assign the ball the set of the ball being green, being blue, and so on, as all of its anti-instantiation criteria. And the ball will exemplify the property of being red iff the state of the ball being red actually obtains, i.e. iff the ball is actually red. And the ball will fail to exemplify the property of being red iff one of the states of the ball being green, the ball being blue, and so on obtains, i.e. iff the ball is colored in some other way than red.

Note that so far we haven't postulated that the exemplification criteria and anti-exemplification criteria of a property for an object must not overlap. Nor have we postulated that at least one of the exemplification criteria or anti-exemplification criteria of a property for an object must actually obtain. Thus, it is possible according to **ECT** that an object both exemplifies and anti-exemplifies a given property and it's possible according to **ECT** that it fails to do either. This is not a bug, it is rather a feature. In this way, we can recover the intuitions of the impossible worlds theory in our framework. If we don't share these intuitions, then we can simply make **ECT** *classical* by additionally postulating the following two conditions for all properties $\Phi = (\Phi^+, \Phi^-)$:

- for all objects x, not both $\Phi^+(x) \cap S_@ \neq \emptyset$ and $\Phi^-(x) \cap S_@ \neq \emptyset$, and
- for all objects x, $\Phi^+(x) \cap S_@ \neq \emptyset$ or $\Phi^-(x) \cap S_@ \neq \emptyset$.

We might call these conditions the *classicality conditions*. They give us a classical conception of hyperintensional properties in the sense that we'll easily be able to show that every object either exemplifies or anti-exemplifies every property. We could make similar adjustments to accommodate other background logics, such negative and positive free-logic, intuitionistic logic, relevant logic and so on.[12] We could also restrict state fusion to only those subsets of the possible sets that are intuitively *compossible* to accommodate the intuition that only possible conditions should play a role in the semantics.[13]

Note also that according to **ECT**, properties are *pairs of functions* and thus individuated as such. We get the following identity criterion for properties

[12] In the last chapter, we've indeed made such an adjustment for the case of negative free logic. Here we don't write down the conditions again, since we're interested in the more general notion of a (hyperintensional) property.

[13] This is indeed what Fine [38] does.

according to **ECT**:

Property Identity in ECT. For all properties $\Phi = (\Phi^+, \Phi^-)$ and $\Psi = (\Psi^+, \Psi^-)$, $\Phi = \Psi$ iff for all $x \in D_\Diamond$ $\Phi^+(x) = \Psi^+(x)$ and $\Phi^-(x) = \Psi^-(x)$.

This criterion allows us to distinguish between necessarily equivalent but intuitively distinct properties, without being too discriminating.

To illustrate let's consider a toy example. Imagine a ball factory that produces balls in one of two colors: blue and red. The factory can eventually produce arbitrarily many balls $b_1, b_2, \ldots$, but so far it has only produced two: b_1 and b_1. In our factory, the color of a ball being produced is decided in the very last production step when the balls are being sprayed either blue or red. Before this production step there is a machine which randomly determines for the balls on the conveyor belt whether they will be painted blue or red. Thus, it is possible for every ball to be painted in either color. In fact, b_1 was painted blue and b_2 was painted red. Finally, the factory is metaphysically possible and thus all the balls that it produces are extended.

If we restrict ourselves to the states of the balls in our toy factory toy example, then we get the following parameters:

- $D_\Diamond = \{b_1, b_2, \ldots\}$
- $D_@ = \{b_1, b_2\}$,
- $S_\Diamond = \{b_i \text{ being blue}, b_i \text{ being red}, b_i \text{ being extended}, b_i \text{ being blue} \circ b_j \text{ being red}, \ldots \mid i, j \in \mathbb{N}\}$,
- $S_@ = \{b_1 \text{ being blue}, b_2 \text{ being red}, b_1 \text{ being extended}, b_2 \text{ being extended}, b_1 \text{ being blue} \circ b_2 \text{ being red}, \ldots\}$.

In this setting, we can define the properties of being red and being blue as the function pairs (being red$^+$, being red$^-$) and (being blue$^-$, being blue$^-$), where these functions are defined by saying for all $b_i \in D_\Diamond$ that:

- being red$^+(b_i) = \{b_i$ being red$\}$,
- being red$^-(b_i) = \{b_i$ being blue$\}$,
- being blue$^+(b_i) = \{b_i$ being blue$\}$, and
- being blue$^-(b_i) = \{b_i$ being red$\}$.

And we can understand the property of being colored as the function pair (being colored$^+$, being colored$^-$), where these functions are defined by saying for all $b_i \in D_\Diamond$ that:

- being colored$^+(b_i) = \{b_i$ being blue, b_i being red, b_i being blue $\circ$ b_i being red$\}$, and

- being colored$^-(b_i) = \emptyset$,

Note that all balls our factory produces are colored and thus the property of being colored has no exact anti-exemplification criteria for our balls. And since we haven't forbidden intuitively non-compossible states to be fused, we have to count the intuitively impossible state of b_i being blue $\circ$ b_i being red as an exact instantiation condition of b_i being colored to satisfy our overdetermining condition.

Moreover, we can understand the property of being extended as the function pair (being extended$^+$, being extended$^-$), where these functions are defined by saying for all $b_i \in D_\Diamond$ that:

- being extended$^+(b_i) = \{b_i$ being extended$\}$, and
- being extended$^-(b_i) = \emptyset$.

Again, since the balls our factory produces are all extended, the property has no exact anti-exemplification criteria for our balls.

Given this setup, we can easily determine using **Property Identity in ECT** that the property of being colored is distinct from the property of being extended: the functions they are defined by give different values for the same arguments. We have found a way of distinguishing the necessarily equivalent but intuitively distinct properties of being colored and being extended. And we have achieved this by what we hold is an intuitively plausible story: properties should be individuated based on what it means to exemplify them. Being colored means to be in a state of having a certain color and being extended means, well, being spatially extended. We could construct similar intuitive toy models for the other property pairs that we mentioned at the outset of this paper, but for reasons of space we shall refrain from doing so.

Note that we can also accomodate the intuition that the properties of being a bachelor and of being an unmarried man are in fact identical. To see this, let's assume that for every object $x \in D_\Diamond$ there are possible states of x being a man, of x being a woman, of x being married, and of x being unmarried in $S_\Diamond$, but *no* distinguished state of x being a bachelor as of yet. In such a setting we may simply *define* the state of being a bachelor as the state of being an unmarried man by saying for all $x \in S_\Diamond$:

$$x \text{ being a bachelor} =_{def.} (x \text{ being a man} \circ x \text{ being unmarried}).$$

This would give intuitive justice to the intuitively plausible idea that being a bachelor means being an unmarried man in virtue of the meaning of "bachelor" being defined by the meaning of "unmarried man" by means of semantic stipulation. In this setting, then, we can plausibly understand the property of being an unmarried man as the pair (unmarried man$^+$,unmarried man$^-$), where these functions are defined by saying for all $x \in D_\Diamond$ that:

- unmarried $\text{man}^+(x) = \{x \text{ being a man} \circ x \text{ being unmarried}\}$, and
- unmarried $\text{man}^-(x) = \{x \text{ being a women}, x \text{ being married}, x \text{ being a woman} \circ x \text{ being married}\}$.

And based on the intuition that being a bachelor *means* being an unmarried man, we can plausibly understand the property of being a bachelor as the pair (bachelor^+,bachelor^-), where these functions are defined by saying for all $x \in D_\Diamond$ that:

- $\text{bachelor}^+(x) = \{x \text{ being a bachelor} = (x \text{ being a man} \circ x \text{ being unmarried})\}$, and
- $\text{bachelor}^-(x) = \{x \text{ being a women}, x \text{ being married}, x \text{ being a women} \circ x \text{ being married}\}$.

Now, under these assumptions, we get in fact that the properties of being a bachelor and of being an unmarried man are identical. Admittedly, this was built into the set-up, as it were, but we've built this into the set-up together with an intuitively plausible background story. The point is merely that we *can* accomodate for the two properties to be identical, we don't *have to*.

Note further that we can also define the Boolean operations **neg**, **disj**, and **conj** on the properties. Inspired by the semantic clauses of Fine [42, 38], which we've discussed in the last chapter, we can define these functions on all properties Φ and Ψ by saying for all $s \in D_\Diamond$ that:

- *Negation*:
  - $\mathbf{neg}(\Phi)^+(x) = \Phi^-(x)$
  - $\mathbf{neg}(\Phi)^-(x) = \Phi^+(x)$
- *Disjunction*:
  - $\mathbf{disj}(\Phi, \Psi)^+(x) = \Phi^+(x) \cup \Psi^+(x) \cup \mathbf{conj}(\Phi, \Psi)^+(x)$
  - $\mathbf{disj}(\Phi, \Psi)^-(x) = \{s \circ t \mid s \in \Phi^-(x), t \in \Psi^-(x)\}$
- *Conjunction*:
  - $\mathbf{conj}(\Phi, \Psi)^+(x) = \{s \circ t \mid s \in \Phi^+(x), t \in \Psi^+(d)\}$
  - $\mathbf{conj}(\Phi, \Psi)^-(x) = \Phi^-(x) \cup \Psi^-(x) \cup \mathbf{conj}(\Phi, \Psi)^-(x)$

It is easily checked that under these operations, the following laws hold:

- $\mathbf{neg}(\mathbf{neg}(\Phi)) = \Phi$
- $\mathbf{disj}(\Phi, \Phi) = \Phi$
- $\mathbf{conj}(\Phi, \Phi) = \Phi$
- $\mathbf{neg}(\mathbf{conj}(\Phi, \Psi) = \mathbf{disj}(\mathbf{neg}(\Phi), \mathbf{neg}(\Phi))$

Moreover, if we postulate further conditions on the properties, such as the classicality conditions, we can show that the properties satisfy further laws, e.g. the laws of a Boolean algebra in the case of the classicality conditions.

But here we are not so interested in the formal properties, as much as in their intuitive applications. Let's go once more into the setting we've described before to show that we can accommodate the intuition that being a bachelor is the same as being an unmarried man. In the setting we've described we can furthermore plausibly postulate that the property of being a man is the pair (man^+,man^-), which is defined by saying for all $x \in D_\Diamond$ that:

- $\text{man}^+(x) = \{x \text{ being a man}\}$, and
- $\text{man}^-(x) = \{x \text{ being a women}\}$.

And similarly, we can plausibly postulate that that the property of being married is the pair (married^+,married^-), which is defined by saying for all $x \in D_\Diamond$ that:

- $\text{married}^+(x) = \{x \text{ being married}\}$, and
- $\text{married}^-(x) = \{x \text{ being unmarried}\}$.

But then we can show that **conj**(being unmarried, being a man)=being an unmarried man=being a bachelor under our stipulations. In other words, we can show being a bachelor is a conjunctive property formed from the two properties of being unmarried and being a man. Thus, we've overcome the limitation of the quasi-syntactic structured properties theory: we can have a conjunctive property being identical to a primitive property.

Finally, we'd like to point to a nice feature of our theory of properties having to do with the theory of hyperintensional propositions. In unpublished work, Fine [46, 47] develops different concepts of hyperintensional propositions in the framework of exact truthmaker semantics. One of the concepts of hyperintensional propositions that Fine discusses is as pairs of sets of possible states, i.e. objects of the form (X, Y) for $X, Y \subseteq S_\Diamond$. The idea is that for such a proposition (X, Y) the members of X are the exact verifiers of the proposition and the members of Y are the exact falsifiers of the proposition in the sense that we've discussed in the last chapter. Thus, Fine simply reifies the semantics that we've discussed in the last chapter and turns it into a theory of propositions. In a sense, what we've done in this chapter is the same for the concept of a property. And indeed, there is a nice connection between Fine's notion of hyperintensional propositions and our properties on the theory **ECT**. First note that Fine discusses various conditions conditions that we might plausibly put on hyperintensional propositions in this sense, among them the condition that for a hyperintensional proposition of the form (X, Y) we want that for all $Z \subseteq X$, we have $\prod Z \in X$ and similarly that $Z \subseteq Y$, we have $\prod Z \in Y$. But now it should be clear that we can canonically interpret properties in the sense of **ECT** as functions that map objects to hyperintensional propositions that satisfy this condition. This is a nice connection between Fine's reification of exact truthmaker semantics to a theory of hyperintensional propositions and our reification of the semantics to a theory of hyperintensional properties. Indeed, we may even understand hyperintesional properties in Fine's sense as a special case of properties, by canonically taking them to be *zero-ary* functions with constant output. This way of treating propositions as a special case of functions, as well as the idea that properties are functions that map objects to propositions is common in Montague semantics for natural language [99] and the full semantics for second order logic [127]. That this connection also holds for our treatment of properties in **ECT** should give some backing to the claim that the exemplification criteria theory of properties is a natural theory of properties in the context of exact truthmaker semantics.

4.5 Conclusion

Let's take stock. In this chapter, we've argued that the go-to hyperintensional theories of properties, the impossible worlds theory and the structured properties theory are not up to the job of giving us a hyperintensional theory of properties which satisfies our intuitive desiderata **Grainedness**, **Ontology**, and **Applications**. In particular, we've argued that the impossible worlds theory fails with regard to **Grainedness** and **Applications**, while the structured properties theory **ASPT** fails with regard to **Ontology** and the quasi-syntactic structure property theory **QSPT** (doubly) fails with regard to **Grainedness**. Motivated by the failures of these theories, we've developed what we hold to be a natural and intuitively motivated theory of hyperintensional properties in the exemplification criteria theory **ECT**.

Our discussion of **ECT** in §4.4 makes it plausible that **ECT** fares quite well with regard to our three desiderata: First, the theory gives us an identity criterion for properties in the form of **Property Identity in ECT** and this criterion allows us to model many intuitively correct claims. In particular, it allows us to overcome the difficulties that **IWT** and **QSPT** have with regard to the individuation of properties: for all of the problematic cases that we've pointed out for these theories, **ECT** as a workable, and so we believe, plausible answer. Second, the theory gives us a clear definition of what a property is: according to **ECT** properties are pairs functions from individuals to exact exemplification and anti-exemplification criteria. Indeed in the §4.4, we've shown that we can even understand properties according to **ECT** more specifically as functions from objects to hyperintensional propositions. Thus, **ECT** fares quite well with regard to the desideratum **Ontology**. And finally, we've already pointed out many times to the various applications of the framework of exact truthmaker semantics and since our theory **ECT** is formulated in this framework in can be put to work in all of these applications. One particular application that is especially pertinent to the goals of this dissertation, namely the explication of the concept of essential properties given by the view Essence Grounded in Haecceities **EGH**, will be discussed in the conclusion of this dissertation. Thus, all in all, the exemplification criteria theory **ECT** appears to pass the test of our three criteria with flying colors. We can arguably regard **ECT** as a good explication of the hyperintensional concept of properties introduced in the introduction of this chapter.

Before we close, we'd like to point out two possible generalizations of the theory we developed in this chapter. First, it would be desirable to extend the theory to a hyperintensional theory of *relations*. This can be achieved without much difficulty by simply defining an n-ary relation to be a pair of n-ary functions from n-tuples of objects to sets of exact exemplification and anti-exemplification criteria. In this chapter we've refrained from doing so for clarity's sake. We hold that the simple case of properties or unary relations illustrates the idea best. Second, it would be desirable to further investigate the algebraic properties of our properties according to **ECT**. We've already pointed out that under natural definitions, some of the laws that hold for properties according to **ECT**. Now it would surely be fruitful to compare the algebraic properties of the properties according to **ECT** with the concrete property algebras described by Bealer [8], Bealer and Mönnich [9], and Menzel [98]. Once we've extended the framework to account for relations and not only properties, it would be especially interesting to consider operations like the hybrid operation $\mathbf{plug_n}$ of Menzel [98] which plugs an *object* $x \in D_\Diamond$ into the n-th *place* of an n-ary relation Σ and turns it into an $(n-1)$-ary relation $\mathbf{plug_n}(x, \Sigma)$ or the operation $\mathbf{\Pi_n}$ which intuitively *universally quantifies* over the n-th place of a relation. But all of this is just subject matter for future research.

Conclusion

Summary

Let's give a brief summary of what we've achieved in this dissertation. In Chapter 1, our introduction, we've set the goals for the dissertation. In particular, by discussing the modal analysis of essential properties **MA** and its explication in the possible worlds framework by means of the possible worlds analysis **PWA** supplemented with the intensional property theory **IPT**, we've come up with two concrete goals: (1) to give a comparable explication to **PWA** for the view that essential properties are grounded in the haecceities of things **EGH**, and (2) to supplement this explication with a comparable property theory to **IPT** which accommodates the hyperintensionality of the concept defined by **EGH**. Then over the core chapters of this dissertation, we've carried out the grunt work for this project. Thereby we've laid the foundations for tackling questions (1) and (2).

In Chapter 2, we've followed an approach to ground that we have ultimately discarded for the purpose of this dissertation: developing an axiomatic theory of conceptualist ground on a predicational approach to the syntax of ground. We've shown that the project of the chapter is well motivated and hopefully it will prove to be fruitful in the future. We have shown many promising results in the chapter, but ultimately the merit of the approach will have to be ascertained by its applicability to other, non-internal philosophical questions. However, in the chapter we've attested that research into axiomatic theories of ground is too young a field of research to find applications to metaphysical problems outside the scope of concrete ground-theoretic issues. Indeed, the chapter is the first proper treatment of axiomatic theories of ground that we're aware of. An especially pressing problem that presented itself is that once we allow for iterated statements of ground in the framework of the chapter, we run into serious difficulties: we face the threat of paradox. For this reason we went looking for other options.

In Chapter 3, we've then turned to operational approaches to the syntax worldly of ground. In this chapter we've extended Fine's semantic framework of exact truthmaker semantics to account for iterated ground in simple predicate languages without quantification. And indeed we could find intuitively plausible clauses for the exact verifiers and falsifiers of equations, atomic sentences, and statements of ground. These are, as far as we're aware, the first clauses that have been given for such formulas in the semantics of ground. We've formulated these clauses with two specific philosophical views in mind: first, the view that grounds ground ground, i.e. the view that grounding facts obtain in virtue of the grounds in the grounding fact; and second, the intuition that atomic statements can only be true if all the terms in them denote existing objects—the core intuition of negative free logic. After giving these clauses, we've surveyed the resulting logic determined by this semantics. We've presented a proposal for a natural deduction system for this logic and proved soundness. Finally, we've sketched how a completeness result may be obtained. However, since the issue of completeness brings us deep into the realm of infinitary logic, we've omitted the details and stayed content with a sound proof theory.

Finally, in Chapter 4, we've developed a hyperintensional property theory that we hold is both natural and will prove fruitful for the aim of this dissertation: we'll argue that the theory can fill the role that **IPT** plays for **PWA** for our analysis **EGH**. This is what we'll argue in this conclusion. In the rest of this conclusion, we will provide a semantic explication of **EGH** in the framework developed in Chapter 3 and we'll subject it to the test of Carnap's four criteria for the quality of an explication.

Explicating Essence Grounded in Haecceities

Now it is time to give our explication of the view that essential properties are properties grounded in the identity of things. For this purpose let us move to the semantic framework of Chapter 3. More specifically, let's assume that we're given a state space $\mathcal{S} = (D_\Diamond, D_@, S_\Diamond, S_@, \prod)$, where $D_\Diamond$ contains all intuitively possible objects, $D_@$ contains all intuitively existing objects, $S_\Diamond$ contains all the intuitively possible states, $D_@$ contains all the intuitively actually obtaining states, and $\prod$ is the intuitive operation of state fusion. It is plausible to assume that $\mathcal{S}$ is indeed a state space, meaning in particular that $S_@$ is closed under (arbirtary) state fusion.

Next, let us regiment our informal meta-language as containing a language of ground $\mathcal{L}$, which contains among its non-logical vocabulary: individual constants for all the objects that we've considered throughout this dissertation, a predicate for every property that we've considered throughout this dissertation, and finally a distinguished *haecceity predicate* $\mathfrak{H}_x$ for every possible object $x \in D_\Diamond$. Thus, for example, $\mathcal{L}$ will contain the individual constants "Socrates" and "Xanthippe;" $\mathcal{L}$ will contain the predicate symbols "being a man," "being married to Xanthippe," "being a member of one's singleton," and “being such that infinitely many prime numbers exists;” and finally $\mathcal{L}$ will contain the haecceity predicates $\mathfrak{H}_{\text{Socrates}}$ for the property of being Socrates and $\mathfrak{H}_{\text{Xanthippe}}$ for the property of being Xanthippe. The properties of being self-identical and being distinct from something, we can express by means of the identity predicate $=$, which is in any case in the logical vocabulary of $\mathcal{L}$. If $x \in D_\Diamond$ is an object, then we'll write $\overline{x}$ for the corresponding individual constant in our language, and if Φ is a property, then we'll write $\overline{\Phi}$ for the corresponding predicate in our language. We may call this language our *essentialist language of ground for $\mathcal{S}$*.[14]

And finally, we assume that we're given an interpretation $\mathcal{I} = (\delta, v^+, v^-)$ for $\mathcal{L}$ in $\mathcal{S}$, which assigns the intended interpretations to all the unproblematic expressions of $\mathcal{L}$. So for example, we can assume that δ(“Socrates”)=Socrates, and δ(“Xanthippe”)=Xanthippe. Or we can assume that Socrates being a man $\in v^+$(“being a man”)(Socrates) or Xanthippe being a woman $\in v^-$(“being a man”)(Xanthippe). As we'll see shortly, the central question will be what v^+ and v^- should assign to the haecceity predicates $\mathfrak{H}_x$. One thing appears to be clear, however. We should get for all $y \neq x \in D_\Diamond$ that $v^+(\mathfrak{H}_x)(y) = \emptyset$ and $v^-(\mathfrak{H}_x)(y) = v^+(\mathfrak{H}_x)(x)$. In words, there is no exact verifier of something other than a given object being that object and an exact falsifier of something being a given object is whatever is an exact verifier of that object being that object.

With this setup in place, the first thing we should do is to properly regiment **EGH** in the framework. Here we have two options, we can understand the “in virtue of” in **EGH** in terms of weak or strict ground. Correspondingly, we get two precisifications of **EGH**:

Essence Weakly Grounded in Haecceities (EWGH). For all properties Φ and for all objects x,

$$\Phi \text{ is an essential property of } x \text{ iff } \mathfrak{H}_x(\overline{x}) \leq \overline{\Phi}(\overline{x})$$

Essence Strictly Grounded in Haecceities (ESGH). For all properties Φ and for all objects x,

$$\Phi \text{ is an essential property of } x \text{ iff } \mathfrak{H}_x(\overline{x}) < \overline{\Phi}(\overline{x})$$

At this point we don't see any reasons for preferring one of the two over the other, the question appears to be a matter of preference. Note, however, that these formulation of the principles are not really full generalized claims, but rather *schemata*. The reason is that we cannot, using the methods of this dissertation, quantify into statements of ground. Moreover, we cannot ascertain the grounds of a claim that some property is essential property of an object, since statements of the form “Φ is an essential property of

[14]In more technical terminology, we assume that $\mathcal{L}$ is *full* with respect to the properties and objects in $\mathcal{S}$: we have predicates and constants for all the properties and objects in $\mathcal{S}$.

x" are not part of our language $\mathcal{L}$. And furthermore, **ESGH** and **ESGH** are merely (metaphysically necessary) equivalences, and this alone doesn't guarantee sameness of grounds to begin with. For these reasons it is useful to *define* predicates of the form "$\overline{\Phi}$ is weakly essential to $\overline{x}$" and "$\overline{\Phi}$ is strictly essential to $\overline{x}$" in our language $\mathcal{L}$ for all properties Φ and all objects x by saying that:

- $\overline{\Phi}$ is weakly essential to $\overline{x} =_{def.} \mathfrak{H}_x(\overline{x}) \leq \overline{\Phi}(\overline{x})$, and
- $\overline{\Phi}$ is strictly essential to $\overline{x} =_{def.} \mathfrak{H}_x(\overline{x}) < \overline{\Phi}(\overline{x})$.

On thing that we can already note at this point is that we can now *show* in our logic that if a property is essential to an object x, then it is an essential property of the object that the property is essential, both according to **EWGH** and **ESGH**. The proof is a simple application of the iteration rules:

$$\frac{\Gamma \leq \varphi}{\Gamma \leq (\Gamma \leq \varphi)} \quad \frac{\Gamma < \varphi}{\Gamma < (\Gamma < \varphi)}$$

Assume, for example, that "being a man is weakly essential to Socrates" is true. This statement is by the above definition simply

"$\mathfrak{H}_{\text{Socrates}}$(Socrates)$\leq$being a man(Socrates),"

and thus we can infer by the first rule (rather its soundness) that

"$\mathfrak{H}_{\text{Socrates}}$(Socrates)$\leq$ ($\mathfrak{H}_{\text{Socrates}}$(Socrates)$\leq$being a man(Socrates))" is true.

But this is just the statement

"($\mathfrak{H}_{\text{Socrates}}$(Socrates)$\leq$being a man(Socrates)) is weakly essential to Socrates."

Hence, we have shown that if it is weakly essential to Socrates that he is a man, then it is weakly essential to Socrates that it is weakly essential to Socrates that he is a man. The argument in the strict case is completely analogous. In other words, we get a principle that we may paraphrase by saying that *essentiality is essential.*

The central question is now: how should v^+ and v^- behave so that we can validate our intuitive paradigmatic examples of essential properties? To tackle this question, first observe that a statement of the form

$$\mathfrak{H}_x(\overline{x}) \leq \overline{\Phi}(\overline{x})$$

is true according to our semantics iff (i) some member of $[\![\mathfrak{H}_x(\overline{x})]\!]^+$ is a member of $S_@$ and (ii) $[\![\mathfrak{H}_x(\overline{x})]\!]^+ \subseteq [\![\overline{\Phi}(\overline{x})]\!]^+$. Similarly, a statement of the form

$$\mathfrak{H}_x(\overline{x}) < \overline{\Phi}(\overline{x})$$

is true according to our semantics iff (i) some member of $[\![\mathfrak{H}_x(\overline{x})]\!]^+$ is a member of $S_@$, (ii) $[\![\mathfrak{H}_x(\overline{x})]\!]^+ \subseteq [\![\overline{\Phi}(\overline{x})]\!]^+$, and (iii) for no $\Gamma \subseteq \mathcal{L}$ is the statement

$$\overline{\Phi}(\overline{x}), \Gamma \leq \mathfrak{H}_x(\overline{x})$$

true. In words, a property Φ is weakly essential to an object x iff some exact verifier of $\mathfrak{H}_x(\overline{x})$ is an actual state and every exact verifier of $\mathfrak{H}_x(\overline{x})$ is an exact verifier of $\overline{\Phi}(\overline{x})$. And analogously Φ is a strictly essential property of an object x iff some exact verifier of $\mathfrak{H}_x(\overline{x})$ is an actual state, every exact verifier of $\mathfrak{H}_x(\overline{x})$ is an exact verifier of $\overline{\Phi}(\overline{x})$, but no exact verifier of $\overline{\Phi}(\overline{x})$ fused with the exact verifier of any other truth is an exact verifier of $\mathfrak{H}_x(\overline{x})$. This gives us a pretty good idea for how v^+ and v^- have to behave so that we can validate our intuitive paradigmatic examples of essential properties in the weak sense: we should make sure that there is some exact verifier of $\mathfrak{H}_x(\overline{x})$ iff $x \in D_@$ and that all of these exact verifiers are exact verifiers of whatever we want to be essential to x. And if we wish to validate our examples in the strict sense we should additionally make sure that no exact verifier of whatever we wish to hold as essential to x can be fused in some way to become a verifier of $\mathfrak{H}_x(\overline{x})$.

Motivated by this observation, we make the following suggestion. Let's assume that for every $x \in D_\Diamond$ there is a unique distinguished state of x being itself$\in S_\Diamond$. We postulate that x being itself $\in S_@$ iff $x \in D_@$. And furthermore, we postulate that for no $X \neq \{\text{x being itself}\} \subseteq S_\Diamond$ is $\prod \Gamma = x$ being itself. Thus, intuitively, we essentially postulate that for every existing object there is an actual atomic state of that object being itself. This state is intuitively the one and only exact verifier of $\mathfrak{H}_x(\overline{x})$—it is for all intents and purposes the haecceity state of x. Correspondingly, we postulate for v^+ and v^- that:

- $v^+(\mathfrak{H}_x)(y) = \begin{cases} \{x \text{ being itself}\} & \text{if } x = y \\ \emptyset & \text{otherwise} \end{cases}$
- $v^-(\mathfrak{H}_x)(y) = \begin{cases} \{\lambda\} & \text{if } x \neq y \\ \emptyset & \text{otherwise} \end{cases}$

Thus, a sentence of the form $\mathfrak{H}_x(c)$ only has an exact falsifier if $\delta(c)$ is x and in that case its only exact falsifier is the haecceity state of x. And a sentence of the form $\mathfrak{H}_x(c)$ only has an exact verifier if $\delta(c)$ is different from x and in that case its only exact verifier is the zero state λ. By the properties we postulated for the haecceity state of x, we can infer that $x \in D_@$ iff some verifier of $\mathfrak{H}_x(\overline{x})$ is in $S_@$, i.e. iff $\mathfrak{H}_x(\overline{x})$ is true. And in case a sentence of the form $\mathfrak{H}_x(c)$ is false, then this is because $\delta(c) \neq x$ and this falsehood is fundamental.

Now, we can say exactly what has to be the case for Φ to be an weakly or strictly essential to an object x. The former is the case iff x exists and the haecceity state of x is the sole exact verifier of $\overline{\Phi}(\overline{x})$ and the latter is the case iff in addition there is some other exact verifier of $\overline{\Phi}(\overline{x})$. This follows immediately from our postulates for $v^+(\mathfrak{H}_x)$ and the properties we have postulated about the haecceity state of x. Thus, if we wish to say that being a man is a weakly essential property of Socrates, then we would have to demand that v^+("being a man")(Socrates)={Socrates being himself}. In other words, we have to demand that the only exact verifier of "Socrates is a men" is the haecceity state of Socrates. Given our initial assumption that Socrates being a man $\in v^+$("being a man")(Socrates), this would mean that Socrates being a man has to be identical to the haecceity state of Socrates. And if we wish to say that being a man is a strictly essential property of Socrates, then we would have to demand that Socrates being himself $\in$ v^+("being a man")(Socrates), but furthermore that there is some other state in v^+("being a man")(Socrates). Since we have assumed that Socrates being a man $\in v^+$("being a man")(Socrates), we could easily achieve this by saying that Socrates is a distinct state from Socrates being a man, while still saying that Socrates being himself $\in v^+$("being a man")(Socrates).

If, in contrast, we wish to say that being a member of his singleton is not an essential property of Socrates, then we would simply have to make sure that the state of Socrates being himself is not an exact verifier of "Socrates is a member of his singleton". And indeed, this is plausible. It would appear that the only exact verifier(s) of "Socrates is a member of his singleton" given that Socrates and his singleton exist, is the haeccaeity state of Socrates' singleton plus perhaps, depending on whether we want to use the strict or the weak concept of essentiality, the state of Socrates being a member of his singleton. In an analogous way, we can make sure that being such that infinitely many prime numbers exist is not an essential property of Socrates: we simply and plausibly assume that Socrates haecceity state is not an exact verifier of this the sentence "Socrates is such that that infinitely many prime numbers exist."

One final complication arises. Remember that in Chapter 3, we postulated that the truth of true equations is fundamental in the sense that their only verifiers, given that the constants involved denote existing objects, is the zero-state λ (compare Definition 3.5.5, p. 130). On the above suggestion, however, this means that being self identical cannot be an essential property of any object, unless the haecceity state of the object is the zero-state. But this is implausible, since the zero state intuitively necessarily exists, and thus if it were the haecceity state of some object, since we know that an object exists iff its haecceity state exists, we would get that the object necessarily exists. But this is only plausible for very few objects. Nevertheless, there is an obvious fix: we simply change the exact verifier clauses for equations to account for our intuition that self identity is an essential property of every existing object. Here are two clauses that will do the job:

- $[\![c_1 = c_2]\!]^+ = \begin{cases} v^+(\mathfrak{H}_{\delta(c_1)})(\delta(c_2)) & \text{if } \delta(c_1) \text{ and } \delta(c_2) \in D_@ \text{ and } \delta(c_1) = \delta(c_2) \\ \emptyset & \text{otherwise} \end{cases}$
- $[\![c_1 = c_2]\!]^+ = \begin{cases} v^+(\mathfrak{H}_{\delta(c_1)})(\delta(c_2)) \cup \{\lambda\} & \text{if } \delta(c_1) \text{ and } \delta(c_2) \in D_@ \text{ and } \delta(c_1) = \delta(c_2) \\ \emptyset & \text{otherwise} \end{cases}$

The first clause works for the weak concept of essential properties and the second one for the strict one.

The exact falsifier clauses for equations, however, should remain unchanged. Remember that the only exact falsifier of a false equation is again the zero fact λ (compare Definition 3.5.5, p. 130). Hence the only verifier of a true inequality statement is again λ. But this has the immediate consequence that a true inequality statement can never be exactly verified by the haecceity state of some object, unless the zero state is identical with the object's haecceity state. But the latter intuitively holds for almost no object, and certainly not for Socrates. Thus, even if Socrates and the Eiffel Tower both exist, being distinct from the Eiffel Tower will not turn out to be an essential property of Socrates—exactly as we want to say.

Putting all of the above together, we arrive at the following, final explications of the concept given by **EGH**:

Weak Exact Verifier Analysis (WEVA). For all properties Φ and all objects x, Φ is an essential property of x iff the haecceity state of x actually obtains and it is the sole verifier of $\overline{\Phi}(\overline{x})$.

Strict Exact Verifier Analysis (SEVA). For all properties Φ and all objects x, Φ is an essential property of x iff the haecceity state of x actually obtains, the haecceity state of x is an exact verifier of $\overline{\Phi}(\overline{x})$, and there is at least one other exact verifier of $\overline{\Phi}(\overline{x})$.

These two explications, together with the exemplification criteria theory of properties **ECT**, are the central result of this dissertation. The question is now, which of the two explications to favor. On merely intuitive grounds, we would prefer **SEVA**, but the choice appears to be merely a question of preference.

How do the explications **WEVA** and **SEVA** together with **ECT** fare with regard to Carnap's four criteria: (1) the *similarity* between the explicatum and the explicandum; (2) the *exactness* of the explicatum; (3) the *fruitfulness* of the explicatum; and (4) the *simplicity* of the explicatum? With regard to (1), we have shown in this chapter that we can validate all of our intuitive paradigmatic examples on both explications. With regard to (2), we hope to have shown in this dissertation that all the concepts that we've used in both **WEVA** and **SEVA** are precisely defined in previously understood terminology. Thus also here both explications seem to do well. With regard to (3) fruitfulness, we have pointed out many times over throughout the dissertation that both the semantic framework of exact truthmaker semantics and especially the exemplification criteria theory of properties have applications all over metaphysics. By explicating essential properties in this framework via both **WEVA** and **SEVA**, we can connect the concept to all of these applications—which will hopefully prove to be fruitful. And with regard to (4) simplicity, we belief that, given we understand all the concepts involved and in particular the concept of exact verification and falsification, both explications are deceptively—at least given all the work that we've put into obtaining them. In any case, **WEVA** and **SEVA** are arguably at least as simple as **PWA**. Thus, the explications appear to do quite well. But ultimately, the analyses will have to stand the test of time: we'll have to see if essentialists really use these explications or whether something better comes along. We've done our part.

Closing Remark

Throughout the dissertation, we've already made many remark about how our results can be extended. But before we close, let's make one final remark about the most pressing way in which the results of this dissertation should be extended. This is by introducing necessity *de re* into our framework. There are two reasons for this. First, we want to be able validate the intuitively plausible inference from $\Gamma \leq \varphi$ or $\Gamma < \varphi$ to $\Box(\bigwedge \Gamma \rightarrow \varphi)$. Essentially, given the results of the thesis, validating this inference will allow us to confirm the principle **Essence Implies Weak Necessity** on the conception of essential properties as properties grounded in the identity of things. And second, we want to be able to compare the modal analysis and the ground-theoretic analysis on equal grounds. And since the modal analysis is is framed in terms of necessity *de re*, this would require us to introduce necessity *de re* into our semantic framework of exact truthmaker semantics. But doing this would require us to say what are the exact verifiers and falsifiers of statements of necessity *de re* and this is a hard question. We conjecture that a lot of exciting research will be carried out in this direction in the future, and we hope to be part of it. But this is work for another day.

Appendix A

Yet Another Puzzle of Ground

A.1 Introduction

This appendix contains my paper "Yet Another Puzzle of Ground," which is forthcoming in *Kriterion*. The published version of the paper can be accessed online under:

http://www.kriterion-journal-of-philosophy.org/kriterion/issues/Permanent/Kriterion-korbmacher-01.pdf

The paper won the *SOPhiA Best Paper Award* at the SOPhiA conference, which took place September 2.–4. at the University of Salzburg.

The paper is a supplement to Chapter 2 and is referenced there. It is included here, since it contains relevant results that go somewhat beyond what is discussed in that chapter and which are of general interest to the subject matter of this dissertation. It can be thought of as a supplement to the chapter.

I should like to thank Albert J. J. Anglberger, Hannes Leitgeb, Thomas Schindler, and Ole Thomassen Hjortland for helpful comments and suggestions on this paper.

A.2 Yet Another Puzzle of Ground

Fine defines ground as "the relation of one truth holding in virtue of others" [44, p. 1].[1] Given this definition, it is natural to think that we should formulate axiomatic first-order theories of ground, which formalize ground by means of a relational *ground predicate* of true sentences. Call such theories *predicational theories of ground*. Predicational theories of ground contrast with *operational theories of ground*, which formalize ground by means of a sentential *ground operator* [24, p. 253–54, 42, p. 46–47]. So far, most theories of ground in the literature are operational theories of ground. But there are at least three theoretical reasons for developing predicational theories of ground:

1. *Quantification*: Predicational theories of ground have greater expressive strength than operational theories of ground. In particular, using a ground predicate, we can formalize ground-theoretic principles involving quantification over truths in a natural way. Take, for example, the intuitively plausible claim that every truth is either fundamental or grounded in some other truths. We can straightforwardly formalize this claim using a ground predicate and first-order quantification over truths, but using a ground operator this is impossible. Without the use of non-classical devices, such as propositional quantification, it is impossible to formalize the nested universal and existential quantification over truths in the principle. Using a ground predicate, in contrast, we can formalize the principle comfortably in the purview of classical first-order logic.

2. *Truth and Modality*: Predicational theories of ground allow us to study ground in the same context as truth and modality. It is generally accepted that truth should be treated as a predicate of sentences, and it has recently been suggested to extend this approach to modality as well [59, 81, 57]. There is an obvious connection between ground and truth, since ground is a relation among truths. But there is also a close connection between ground and modality, since ground is usually assumed to imply necessary consequence: if a truth holds in virtue of some other truths, then the former truth should be a necessary consequence of the latter truths [42, p. 38–39]. Both of these connections are most naturally studied using predicational theories of ground: by combining predicational theories of ground with predicational theories of truth and modality.

3. *Models*: Predicational theories of ground allow us to discover and to

[1] For (opinionated) introductions to ground, see [28, 42]. For an overview of the recent literature on ground, see [21, 133, 118].

study models of ground using classic model-theoretic methods. It is currently an open problem to provide a semantics for the *impure logic of ground* developed by Fine [42, p. 58–71]. This logic is formulated using a ground operator, but once we translate it into a predicational theory of ground and show its consistency, we can rely on model-theoretic theorems to establish the existence of first-order models. Once we know that such models exist, we can study them using methods of model theory. This should provide us with new insights into the semantics of the impure logic of ground.

But predicational theories of ground face a paradox of self-reference, similar to the well-known paradoxes of self-reference that arise in predicational theories of truth and modality. In this paper, I shall prove this point for predicational theories of *partial ground* in particular. This is the relation of one truth holding partially in virtue of another truth—the relation of one truth "helping" to ground another truth [42, p. 50]. I show that any predicational theory of partial ground that extends a standard theory of syntax and that proves some commonly accepted principles for partial ground is inconsistent. Fine [43] and Krämer [73] present puzzles about the irreflexivity of partial ground: the principle that no truth partly grounds itself. They show that certain intuitively plausible principles of logic and metaphysics lead to counterexamples to the irreflexivity of ground. I add yet another puzzle of ground to the mix. The new puzzle does not mention the irreflexivity of ground or metaphysical principles unrelated to ground, thus it is genuinely different from the previously known paradoxes.

To formulate a predicational theory of partial ground, we first need a theory of syntax that allows us to talk about sentences.[2] It is well-known that we can develop such a theory in any sufficiently strong background theory, like Robinson arithmetic for example. For the present purpose, however, our concrete choice of background theory does not matter. All that matters is that our background theory Θ satisfies the following three minimal syntax conditions:[3] The first condition is that Θ proves that we have a unique name $\ulcorner\varphi\urcorner$ for every sentence φ in the sense that for all sentences φ and ψ, $\Theta \vdash \ulcorner\varphi\urcorner = \ulcorner\psi\urcorner$ only if $\varphi = \psi$. The second condition is that Θ proves that we have a function symbol $\underline{\vee}$ that represents the syntactic operation $\vee$ of disjunction in the sense that for all sentences φ and ψ, $\Theta \vdash \ulcorner\varphi\urcorner \,\underline{\vee}\, \ulcorner\psi\urcorner = \ulcorner\varphi \vee \psi\urcorner$. And the third condition is that Θ proves the diagonal lemma. Informally, this lemma states that for every condition on sentences there is a sentence that is provably equivalent to the condition holding of itself. More precisely, if $\varphi(x)$ is a formula with exactly one free variable, then there exists a sentence

[2] A sentence is a formula without any free variables.

[3] A theory is a set of formulas that is closed under derivability: a set of formulas Θ is a theory iff (if and only if) for all formulas φ, if $\Theta \vdash \varphi$, then $\varphi \in \Theta$.

δ such that $\Theta \vdash \delta \leftrightarrow \varphi(\ulcorner\delta\urcorner)$. Note that any standard background theory of syntax, such as Robinson arithmetic, satisfies all three of our minimal syntax conditions.

Next, we need a way of representing partial ground. For this purpose, we use the relational predicate $x \lhd y$. For sentences φ and ψ, we informally read the atomic formula $\ulcorner\varphi\urcorner \lhd \ulcorner\psi\urcorner$ as saying that the truth of φ partially grounds the truth of ψ. For a negated atomic formula of the form $\neg(\ulcorner\varphi\urcorner \lhd \ulcorner\psi\urcorner)$ we also write $\ulcorner\varphi\urcorner \ntriangleleft \ulcorner\psi\urcorner$, which we correspondingly read as saying that the truth of φ does not even partially ground the truth of ψ.

Philosophers have laid down various principles for partial ground [cf. 121, 42, 44], but it is already sufficient for a predicational theory of partial ground to be inconsistent that it proves two widely accepted principles. Let Θ now be a predicational theory of partial ground that satisfies the minimal syntax conditions. The first of our two principles follows directly from partial ground being a relation of true sentences: If the truth of one sentence partially grounds the truth of another, then both sentences should be true. This principle is known as the "factivity of ground" and is generally accepted in the literature on ground [43, p. 100, 15, § 3].[4] We get the condition on Θ that for all sentences φ and ψ:

(Fact$_L$): $\Theta \vdash \ulcorner\varphi\urcorner \lhd \ulcorner\psi\urcorner \rightarrow \varphi$
(Fact$_R$): $\Theta \vdash \ulcorner\varphi\urcorner \lhd \ulcorner\psi\urcorner \rightarrow \psi$

The second principle concerns the interaction of partial ground and disjunction: Given that partial ground is the relation of one truth holding partially in virtue of another, if a disjunction is true, then its truth should be partially grounded in each of its true disjuncts. Also this principle is generally accepted in the literature on ground [43, p. 101, 121, p. 117]. From this, we get the condition on Θ that for all sentences φ and ψ:

($\vee_1$): $\Theta \vdash \varphi \rightarrow \ulcorner\varphi\urcorner \lhd \ulcorner\varphi \vee \psi\urcorner$
($\vee_2$): $\Theta \vdash \psi \rightarrow \ulcorner\psi\urcorner \lhd \ulcorner\varphi \vee \psi\urcorner$

The minimal syntax conditions and the conditions concerning partial ground all may seem fairly uncontroversial when viewed individually. So it may be somewhat surprising to learn that there can be no consistent predicational theory of partial ground that satisfies all of them:

Theorem (Inconsistency Theorem)**.** *Any theory Θ that satisfies the minimal syntax conditions, (Fact$_{L/R}$), and ($\vee_{1/2}$) is inconsistent.*

[4] There are notions of ground in the literature that violate the factivity of ground [cf. 42, p. 48–50]. According to such *non-factive* notions, ground is a relation on sentences regardless of their truth value. Although non-factive notions of ground make for an interesting theoretical possibility, in this paper we shall deal only with the standard *factive* notion of ground, which satisfies the factivity of ground.

Proof. Let $\varphi(x)$ be the formula $x \ntriangleleft x \underline{\vee} x$. By the diagonal lemma, there is a sentence δ such that $\Theta \vdash \delta \leftrightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta\urcorner \underline{\vee} \ulcorner\delta\urcorner$. Intuitively, this is a sentence which "says of itself" that it does not partially ground its own disjunction. By the second minimality condition, we have that $\Theta \vdash \ulcorner\delta\urcorner \underline{\vee} \ulcorner\delta\urcorner = \ulcorner\delta \vee \delta\urcorner$. From this and $\Theta \vdash \delta \leftrightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta\urcorner \underline{\vee} \ulcorner\delta\urcorner$, we get that $\Theta \vdash \delta \leftrightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner$ by the substitutivity of identicals. This splits up into the following two conditions:

(a) $\Theta \vdash \delta \rightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner$

(b) $\Theta \vdash \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner \rightarrow \delta$

We get finally the following argument:

1. $\Theta \vdash (\ulcorner\delta\urcorner \triangleleft \ulcorner\delta \vee \delta\urcorner \rightarrow \delta) \rightarrow ((\delta \rightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner) \rightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner)$ (Tautology†)
2. $\Theta \vdash \ulcorner\delta\urcorner \triangleleft \ulcorner\delta \vee \delta\urcorner \rightarrow \delta$ (Fact$_L$)
3. $\Theta \vdash (\delta \rightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner) \rightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner$ (1, 2: MP)
4. $\Theta \vdash \delta \rightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner$ (a)
5. $\Theta \vdash \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner$ (3, 4: MP)
6. $\Theta \vdash \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner \rightarrow \delta$ (b)
7. $\Theta \vdash \delta$ (5, 6: MP)
8. $\Theta \vdash \delta \rightarrow \ulcorner\delta\urcorner \triangleleft \ulcorner\delta \vee \delta\urcorner$ ($\vee_1$)
9. $\Theta \vdash \ulcorner\delta\urcorner \triangleleft \ulcorner\delta \vee \delta\urcorner$ (7,8: MP)
10. $\Theta \vdash \bot$ (5,9: $\bot$)

($\dagger$) : Note that every sentence of the form $(\varphi \rightarrow \psi) \rightarrow ((\psi \rightarrow \neg\varphi) \rightarrow \neg\varphi)$ is a classical tautology and that theories prove all classical tautologies. □

The inconsistency theorem is very similar to Tarski's theorem about predicational theories of truth [131] and Montague's theorem about predicational theories of modality [101] in that it is, essentially, a paradox of self-reference. From a technical perspective, it should in fact not be surprising that we get such a theorem after all: Combining self-reference via the diagonal lemma with principles like (Fact$_{L/R}$) that allow us to push a sentence outside the scope of a predicate and principles like ($\vee_{1/2}$) that allow us to push a sentence into the scope of a predicate is recipe for disaster.[5] But from a philosophical perspective, there is a lesson to be learned: We already know that we cannot

[5]It should be clear at this point that not much depends on the concrete condition ($\vee_{1/2}$)—the paradox is not a paradox of disjunction. All that matters is that our predicational theory of partial ground proves a principle to the effect that any true sentence partially grounds some other sentence. We could give the following variant of the inconsistency theorem: If Θ satisfies (Fact$_{L/R}$) and either $\Theta \vdash \varphi \rightarrow \exists x(\ulcorner\varphi\urcorner \triangleleft x)$ or $\Theta \vdash \varphi \rightarrow \exists x(x \triangleleft \ulcorner\varphi\urcorner)$, then Θ is inconsistent. I leave the details of the proof to the interested reader.

understand ground simply in terms of truth and modality [42], but ground behaves syntactically too much like a combination of truth and modality to escape inconsistency when paired with self-reference.

Three natural ways in which we could try to block the inconsistency theorem suggest themselves: First, we could try to rule out self-referential sentences of ground like the one used in the proof of the inconsistency theorem. Second, we could try to restrict the principles of partial ground used in the proof of the inconsistency theorem. And third, we could try to formulate a non-standard logic of ground that does not sanction the logical principles used in the proof of the inconsistency theorem. The analogy between the inconsistency theorem and the theorems of Tarski and Montague suggests a terminology for these approaches. Analogously to predicational theories of truth [58] and predicational theories of modality [57], we get: *typed theories of partial ground*, which avoid paradox by putting type-restrictions on the relation of partial ground, effectively ruling out self-referential sentences like the one in the proof; *untyped theories of partial ground*, which avoid paradox by restricting the principles of partial ground; and finally *non-classical theories of partial ground*, which avoid paradox (or: triviality) by abandoning classical logic in favor of alternative logics.

Untyped theories of partial ground are particularly appealing, because considerations of ground are already part of intuitively appealing approach to untyped theories of truth. On an influential view about predicational theories of truth, self-referential sentences are *ungrounded* and this is the reason some self-referential sentences lead to inconsistency [77, 82]. This leads to the idea that we should restrict the principles of truth to their grounded instances.[6] Carrying this idea from theories of truth over to predicational theories of ground, we arrive at the condition that the principles of partial ground apply if and only if the truths involved are themselves grounded. There is a straightforward way of formulating the desired restriction on the principles of partial ground already in the language of partial ground. We can express that a sentence φ is grounded by the formula $\exists x(x \lhd \ulcorner\varphi\urcorner)$ and we can express that a sentence φ is ungrounded by the formula $\neg\exists x(x \lhd \ulcorner\varphi\urcorner)$. The desired restriction on $(\vee_{1/2})$ then amounts to saying that for all predicational theories of ground Θ and for all sentences φ and ψ:

$$(\vee_1^*): \quad \Theta \vdash \exists x(x \lhd \ulcorner\varphi\urcorner) \leftrightarrow \ulcorner\varphi\urcorner \lhd \ulcorner\varphi \vee \psi\urcorner$$
$$(\vee_2^*): \quad \Theta \vdash \exists x(x \lhd \ulcorner\psi\urcorner) \leftrightarrow \ulcorner\psi\urcorner \lhd \ulcorner\varphi \vee \psi\urcorner$$

Every predicational theory of partial ground that satisfies the minimal syn-

[6] The concept of ground used in the context of theories of truth is not exactly the same as the concept of ground discussed in this paper. For example, the notion of dependence defined by Leitgeb [82] is reflexive, whereas (partial) ground is standardly taken to be irreflexive. The point here is that there is a striking analogy between the two concepts and that ideas that work for the one may very well work for the other.

tax conditions, ($\text{Fact}_{L/R}$), and the new conditions ($\vee^*_{1/2}$), proves that the paradoxical sentence in the proof of the theorem is ungrounded:

Observation. *Let Θ be a predicational theory of ground that satisfies the minimal syntax conditions, ($\mathit{Fact}_{L/R}$), and ($\vee^*_{1/2}$). By the diagonal lemma and the same reasoning as in the proof of the theorem, we get a sentence δ such that:*

$$\Theta \vdash \delta \leftrightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner.$$

But we can show that:

$$\Theta \vdash \neg\exists x(x \triangleleft \ulcorner\delta\urcorner).$$

Proof. By applying ($\vee^*_1$) to δ, we get that:

$$\Theta \vdash \exists x(x \triangleleft \ulcorner\delta\urcorner) \leftrightarrow \ulcorner\delta\urcorner \triangleleft \ulcorner\delta \vee \delta\urcorner$$

We only need the "left-to-right direction" of this biconditional for our proof, which we can obtain via $\leftrightarrow$-Elimination:

$$\Theta \vdash \exists x(x \triangleleft \ulcorner\delta\urcorner) \rightarrow \ulcorner\delta\urcorner \triangleleft \ulcorner\delta \vee \delta\urcorner$$

Starting from there, we get the following argument:

1. $\Theta \vdash \exists x(x \triangleleft \ulcorner\delta\urcorner) \rightarrow \ulcorner\delta\urcorner \triangleleft \ulcorner\delta \vee \delta\urcorner$
2. $\Theta \vdash \ulcorner\delta\urcorner \triangleleft \ulcorner\delta \vee \delta\urcorner \rightarrow \delta$ (Fact_L)
3. $\Theta \vdash \exists x(x \triangleleft \ulcorner\delta\urcorner) \rightarrow \delta$ (1,2: MP)
4. $\Theta \vdash \delta \leftrightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner$ (Diagonal Lemma)
5. $\Theta \vdash \exists x(x \triangleleft \ulcorner\delta\urcorner) \rightarrow \ulcorner\delta\urcorner \ntriangleleft \ulcorner\delta \vee \delta\urcorner$ (3, 4: $\leftrightarrow$-Elim)
6. $\Theta \vdash \exists x(x \triangleleft \ulcorner\delta\urcorner) \rightarrow \bot$ (1, 5: $\bot$-Intro)
7. $\Theta \vdash \neg\exists x(x \triangleleft \ulcorner\delta\urcorner)$ (6: $\neg$-Intro)

□

This result should make us optimistic about the prospects for an untyped theory of partial ground. Moreover, we could add such an untyped theory of partial ground "on top" of untyped predicational theories of truth and modality. I conjecture that an interesting, *consistent*, untyped predicational theory of ground can be developed in this way.

Bibliography

[1] A. R. Anderson and N. D. Belnap. "First Degree Entailments". In: *Journal of Symbolic Logic* 36.3 (1971), pp. 520–521.

[2] Richard B. Angell. "Deducibility, Entailment and Analytic Containment". In: *Directions in Relevant Logic*. Ed. by Jean Norman and Richard Sylvan. Kluwer Academic Publishers, 1989, pp. 119–143.

[3] Ralf M. Bader. "Towards a Hyperintensional Theory of Intrinsicality". In: *Journal of Philosophy* 110.10 (2013), pp. 525–63.

[4] John Robert Baker. "Essentialism and the Modal Semantics of J. Hintikka". In: *Notre Dame Journal of Formal Logic* 19.1 (1978), pp. 81–91.

[5] Roberta Ballarin. "Modern Origins of Modal Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2014. 2014.

[6] Arnould Bayart. "Correction de la Logique Modale du Premier Et du Second Ordre S5". In: *Logique Et Analyse* 1.1 (1958), pp. 28–45.

[7] Arnould Bayart. "Quasi-Adéquation de la Logique Modale du Second Ordre S5 Et Adéquation de la Logique Modale du Premier Ordre S5 [Quasi-Completeness of Second-Order S5 Modal Logic and Completeness of First-Order S5 Modal Logic]". In: *Logique Et Analyse* 2.6 (1959), pp. 99–121.

[8] George Bealer. "Theories of Properties, Relations, and Propositions". In: *Journal of Philosophy* 76.11 (1979), pp. 634–48.

[9] George Bealer and Uwe Mönnich. "Property Theories". In: *Handbook of Philosophical Logic*. Ed. by Dov Gabbay and Franz Guenthner. Kluwer, 1989, pp. 133–251.

[10] John L. Bell. "Infinitary Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2012. 2012.

[11] Karen Bennett. "By Our Bootstraps". In: *Philosophical Perspectives* 25.1 (2011), pp. 27–41.

[12] George Berkeley. *A Treatise on the Principles of Human Knowledge.* Aaron Rhames, 1710.

[13] Francesco Berto. *Existence as a Real Property.* Synthèse Library, Springer, 2012.

[14] Francesco Berto. "Impossible Worlds". In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Winter 2013. 2013.

[15] Ricki Bliss and Kelly Trogdon. "Metaphysical Grounding". In: *The Stanford Encyclopedia of Philosophy.* Ed. by Edward N. Zalta. Winter 2014. 2014.

[16] George Boolos et al. *Computability and Logic.* Cambridge University Press, 2007.

[17] Rudolf Carnap. *Logical Foundations of Probability.* Chicago University of Chicago Press, 1962.

[18] Rudolf Carnap. *Meaning and Necessity.* University of Chicago Press, 1947.

[19] Rudolf Carnap. "Modalities and Quantification". In: *Journal of Symbolic Logic* 11.2 (1946), pp. 33–64.

[20] Rudolf Carnap. "The Elimination of Metaphysics Through Logical Analysis of Language". In: *Erkenntnis* (1932), pp. 60–81.

[21] Michael Clark and David Liggins. "Recent Work on Grounding". In: *Analysis* 72.4 (2012), pp. 812–23.

[22] Fabrice Correia. "(Finean) Essence and (Priorean) Modality". In: *Dialectica* 61.1 (2007), pp. 63–84.

[23] Fabrice Correia. "Generic Essence, Objectual Essence, and Modality". In: *Noûs* 40.4 (2006), pp. 753–67.

[24] Fabrice Correia. "Grounding and Truth-Functions". In: *Logique et Analyse* 53.211 (2010), pp. 251–79.

[25] Fabrice Correia. "Logical Grounds". In: *Review of Symbolic Logic* 7.1 (2013), pp. 31–59.

[26] Fabrice Correia. "On the Logic of Factual Equivalence". In: *Review of Symbolic Logic* (forthcoming), pp. 1–20.

[27] Fabrice Correia. "Semantics for Analytic Containment". In: *Studia Logica* 77.1 (2004), pp. 87–104.

[28] Fabrice Correia and Benjamin Schnieder. "Grounding. An Opiniated Introduction". In: *Metaphysical Grounding: Understanding the Structure of Reality.* Ed. by Fabrice Correia and Benjamin Schnieder. Cambridge University Press, 2012, pp. 1–36.

[29] M. J. Cresswell. "Hyperintensional Logic". In: *Studia Logica* 34.1 (1975), pp. 25–38.

[30] Shamik Dasgupta. “The Possibility of Physicalism”. In: *Journal of Philosophy* 111.9/10 (2014), pp. 557–592.

[31] Louis deRosset. “Better Semantics for the Pure Logic of Ground”. In: (2015). forthcoming in *Analytic Philosophy*.

[32] Louis deRosset. “Grounding Explanations”. In: *Philosophers' Imprint* 13.7 (2013).

[33] Louis Derosset. “On Weak Ground”. In: *Review of Symbolic Logic* 7.4 (2014), pp. 713–44.

[34] J. Michael Dunn. “Relevant Predication 3: Essential Properties”. In: *Truth or Consequences*. Ed. by J. Michael Dunn and Anil Gupta. Kluwer, 1990, pp. 77–95.

[35] Solomon Feferman. “Reflecting on Incompleteness”. In: *Journal of Symbolic Logic* 56.1 (1991), pp. 1–49.

[36] Kit Fine. “A Survery of Truthmaker Semantics”. In: (2015). Manuscript published on academia.edu. `https://www.academia.edu/10908756/survey_of_truthmaker_semantics`.

[37] Kit Fine. “Angellic Content”. In: *Journal of Philosophical Logic* (forthcoming), pp. 1–28.

[38] Kit Fine. “Counterfactuals Without Possible Worlds”. In: *Journal of Philosophy* 109.3 (2012), pp. 221–46.

[39] Kit Fine. “Essence and Modality”. In: *Philosophical Perspectives* 8 (1994), pp. 1–16.

[40] Kit Fine. “First-Order Modal Theories I–Sets”. In: *Noûs* 15.2 (1981), pp. 177–205.

[41] Kit Fine. “First-Order Modal Theories III — Facts”. In: *Synthese* 53.1 (1982), pp. 43–122.

[42] Kit Fine. “Guide to Ground”. In: *Metaphysical Grounding: Understanding the Structure of Reality*. Ed. by Fabrice Correia and Benjamin Schnieder. Cambridge University Press, 2012, pp. 37–80.

[43] Kit Fine. “Some Puzzles of Ground”. In: *Notre Dame Journal of Formal Logic* 51.1 (2010), pp. 97–118.

[44] Kit Fine. “The Pure Logic of Ground”. In: *Review of Symbolic Logic* 5.1 (2012), pp. 1–25.

[45] Kit Fine. “Truth-Maker Semantics for Intuitionistic Logic”. In: *Journal of Philosophical Logic* 43.2-3 (2014), pp. 549–77.

[46] Kit Fine. “Truthconditional Content - Part I”. In: (2015). Manuscript published on academia.edu. `https://www.academia.edu/9186494/Truthconditional_Content_-_Part_I`.

[47] Kit Fine. "Truthconditional Content - Part II". In: (2015). Manuscript published on academia.edu. `https://www.academia.edu/9206919/Truthconditional_Content_-_Part_II`.

[48] Kit Fine. "Unified Foundations for Essence and Ground". In: *Journal of the American Philosophical Association* 1.2 (2015), pp. 296–311.

[49] Kit Fine. "Varieties of Necessity". In: *Conceivability and Possibility*. Ed. by Tamar Szabo Gendler and John Hawthorne. Oxford University Press, 2002, pp. 253–81.

[50] Kai von Fintel and Irene Heim. *Intensional Semantics*. Available at `http://web.mit.edu/fintel/fintel-heim-intensional.pdf`. 2011.

[51] Melvin Fitting and Richard L. Mendelsohn. *First-Order Modal Logic*. Kluwer Academic Publishers, 1998.

[52] Dagfinn Føllesdal. "Quine on Modality". In: *The Cambridge Companion to Quine*. Ed. by Roger F. Gibson. Cambridge University Press, 2004, pp. 200–13.

[53] Bas C. van Fraassen. "Facts and Tautological Entailment". In: *The Journal of Philosophy* 66.15 (1969), pp. 477–87.

[54] Gottlob Frege. "Über Sinn Und Bedeutung". In: *Zeitschrift für Philosophie Und Philosophische Kritik* 100.1 (1892), pp. 25–50.

[55] Robert Goldblatt. "Mathematical Modal Logic: a View of its Evolution. in Modalities in the Twentieth Century". In: *Handbook of the History of Logic*. Ed. by Dov M. Gabbay and John Woods. Vol. 7. Elsevier, 2006, pp. 1–98.

[56] Judy Green. "Consistency Properties for finite Quantifier Languages". In: *Infinitary Logic: In Memoriam Carol Karp*. Ed. by David W. Kueker. Springer, 1975, pp. 73–123.

[57] V. Halbach and P. Welch. "Necessities and Necessary Truths: A Prolegomenon to the Use of Modal Logic in the Analysis of Intensional Notions". In: *Mind* 118.469 (2009), pp. 71–100.

[58] Volker Halbach. *Axiomatic Theories of Truth*. Cambridge University Press, 2011.

[59] Volker Halbach, Hannes Leitgeb, and Philip Welch. "Possible-Worlds Semantics for Modal Notions Conceived as Predicates". In: *Journal of Philosophical Logic* 32.2 (2003), pp. 179–223.

[60] Jaakko Hintikka. "Modality and Quantification". In: *Theoria* 27.3 (1961), pp. 119–128.

[61] Leon Horsten. "Impredicative Identity Criteria". In: *Philosophy and Phenomenological Research* 80.2 (2010), pp. 411–39.

[62] G. E. Hughes and Max Cresswell. *A New Introduction to Modal Logic*. Routledge, 1996.

[63] Mark Jago. "Recent Work in Relevant Logic". In: *Analysis* 73.3 (2013), pp. 526–41.

[64] Mark Jago. *The Impossible: An Essay on Hyperintensionality*. Oxford University Press, 2014.

[65] Thomas Jech. *Set Theory*. Springer, 2002.

[66] C. S. Jenkins. "Is Metaphysical Dependence Irreflexive?" In: *The Monist* 94.2 (2011), pp. 267–276.

[67] Stig Kanger. *Provability in Logic*. Stockholm, Almqvist & Wiksell, 1957, p. 376.

[68] Carol Karp. "Nonaxiomatizability Results for Infinitary Systems". In: *Journal of Symbolic Logic* 32.3 (1967), pp. 367–84.

[69] Carol Ruth Karp. *Languages with Expressions of Infinite Length*. North-Holland, 1964.

[70] Jeffrey C. King. *The Nature and Structure of Content*. Oxford University Press, 2007.

[71] William Kneale. "Modality De Dicto and De Re". In: *Logic, Methodology and Philosophy of Science*. Ed. by Ernest Nagel, Patrick Suppes, and Alfred Tarski. Stanford University Press, 1960, pp. 622–33.

[72] Johannes Korbmacher. "Yet Another Puzzle of Ground". In: *Kriterion – Journal of Philosophy* 29.2 (2015), pp. 1–10.

[73] Stephan Krämer. "A Simpler Puzzle of Ground". In: *Thought: A Journal of Philosophy* 2.2 (2013), pp. 85–89.

[74] Saul A. Kripke. "A Completeness Theorem in Modal Logic". In: *Journal of Symbolic Logic* 24.1 (1959), pp. 1–14.

[75] Saul A. Kripke. "Identity and Necessity". In: *Identity and Individuation*. Ed. by Milton K. Munitz. New York University Press, 1971, pp. 135–64.

[76] Saul A. Kripke. *Naming and Necessity*. Harvard University Press, 1980.

[77] Saul A. Kripke. "Outline of a Theory of Truth". In: *Journal of Philosophy* 72.19 (1975), pp. 690–716.

[78] Saul A. Kripke. "Semantical Considerations on Modal Logic". In: *Acta Philosophica Fennica* 16.1963 (1963), pp. 83–94.

[79] Karel Lambert. *Free Logic: Selected Essays*. New Yorkcambridge University Press, 2003, pp. 413–419.

[80] Hannes Leitgeb. "Criteria of Identity: Strong and Wrong". In: *British Journal for the Philosophy of Science* 64.1 (2013), pp. 61–68.

[81] Hannes Leitgeb. "Towards a Logic Of Type-Free Modality and Truth". In: *Logic Colloquium 2005: Proceedings of the Annual European Summer Meeting of the Association for Symbolic Logic, Held in Athens, Greece, July 28-August 3, 2005*. Ed. by Costas Dimitracopoulos. Cambridge University Press, 2008, pp. 28–68.

[82] Hannes Leitgeb. "What Truth Depends On". In: *Journal of Philosophical Logic* 34.2 (2005), pp. 155–92.

[83] David K. Lewis. *On the Plurality of Worlds*. Blackwell, 1986.

[84] Jon Erling Litland. "Grounding, Explanation, and the Limit of Internality". In: (2015). forthcoming in *Philosophical Review*.

[85] Jon Erling Litland. "Grounding Grounding". In: *Oxford Studies in Metaphysics* (forthcoming).

[86] Jon Erling Litland. "On Some Counterexamples to the Transitivity of Grounding". In: *Essays in Philosophy* 14.1 (2013), pp. 19–32.

[87] Jon Erling Litland. "The Pure Logic of Iterated Ground—Full Ground". In: (2014). Unpublished manuscript. *Academia.edu*, `https://www.academia.edu/9100265/Pure_Logic_of_Iterated_Ground_-_Full_Ground` (accessed May 13, 2015).

[88] Jon Erling Litland. "Topics in Philosophical Logic". Doctoral Dissertation, Harvard University. *DASH: Digital Access to Scholarship at Harvard*, `http://dash.harvard.edu/handle/1/9527318` (accessed May 13, 2015). 2012.

[89] Ruth Barcan Marcus. "A Backwards Look at Quine's Animadversions on Modalities". In: *Perspectives on Quine*. Ed. by Robert Barrett and Richard J. Gibson. Blackwell, 1990, pp. 230–43.

[90] Ruth Barcan Marcus. "Essential Attribution". In: *Journal of Philosophy* 68.7 (1971), pp. 187–202.

[91] Ruth Barcan Marcus. "Essentialism in Modal Logic". In: *Noûs* 1.1 (1967), pp. 91–96.

[92] Edwin Mares. "Relevance Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2014. 2014.

[93] Thomas J. McKay. "Essentialism in Quantified Modal Logic". In: *Journal of Philosophical Logic* 4.4 (1975), pp. 423–38.

[94] Alan McMichael and Ed Zalta. "An Alternative Theory of Nonexistent Objects". In: *Journal of Philosophical Logic* 9.3 (1980), pp. 297–313.

[95] Alexius Meinong. "On the Theory of Objects (Translation of 'Über Gegenstandstheorie', 1904)". In: *Realism and the Background of Phenomenology*. Ed. by Roderick Chisholm. Reprint 1960. Free Press, 1904, pp. 76–117.

[96] Christopher Menzel. "Actualism". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2015. 2015.

[97] Christopher Menzel. "Possible Worlds". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2015. 2015.

[98] Christopher Menzel. "The Proper Treatment of Predication in Fine-Grained Intensional Logic". In: *Philosophical Perspectives* 7 (1993), pp. 61–87.

[99] Richard Montague. *Formal Philosophy; Selected Papers of Richard Montague*. New Haven,Yale University Press, 1974.

[100] Richard Montague. "Logical Necessity, Physical Necessity, Ethics, and Quantifiers". In: *Inquiry* 3.1-4 (1960). presented in 1955, pp. 259–69.

[101] Richard Montague. "Syntactical Treatment of Modality, with Corollaries on Reflection Principles and Finite Axiomatizability". In: *Acta Philosophica Fennica* 16 (1963), pp. 153–67.

[102] Michael Nelson. "Existence". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2012. 2012.

[103] Daniel Nolan. "Hyperintensional Metaphysics". In: *Philosophical Studies* 171.1 (2014), pp. 149–160.

[104] John Nolt. "Free Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2014. 2014.

[105] Alex Oliver and Timothy Smiley. "Multigrade Predicates". In: *Mind* 113.452 (2004), pp. 609–81.

[106] Willard van Orman Quine. *Ontological Relativity and Other Essays*. Columbia University Press, 1996.

[107] Charles Parsons. *Mathematical Thought and its Objects*. Cambridge University Press, 2008.

[108] Charles Parsons. "Sets and Modality". In: *Mathematics in Philosophy*. Ed. by Charles Parsons. Cornell University Press, 1983, pp. 298–341.

[109] Terence Parsons. "Grades of Essentialism in Quantified Modal Logic". In: *Noûs* 1.2 (1967), pp. 181–191.

[110] Alvin Plantinga. *The Nature of Necessity*. Oxford University Press, 1974.

[111] Wolfram Pohlers. *Proof Theory. The First Step Into Impredicativity.* Springer, 2009.

[112] Graham Priest. "Sylvan's Box: A Short Story and Ten Morals". In: *Notre Dame Journal of Formal Logic* 38.4 (1997), pp. 573–82.

[113] Graham Priest. *Towards Non-Being: The Logic and Metaphysics of Intentionality.* Vol. 14. 1. Oxford University Press, 2005, p. 116.

[114] A. N. Prior. *Time and Modality.* Greenwood Press, 1955, p. 266.

[115] W.V.O. Quine. *Mathematical Logic.* Harvard University Press, 1951.

[116] W.V.O. Quine. "Three Grades of Modal Involvement". In: *Journal of Symbolic Logic*. 2. North-Holland Publishing Co., 1953, pp. 168–169.

[117] W.V.O. Quine. *Word and Object.* The MIT Press, 1960.

[118] Michael J. Raven. "Ground". In: *Philosophy Compass* 10.5 (2015), pp. 322–33.

[119] Michael J. Raven. "Is Ground a Strict Partial Order?" In: *American Philosophical Quarterly* 50.2 (2013), pp. 191–99.

[120] David Ripley. "Structures and Circumstances: Two Ways to Fine-Grain Propositions". In: *Synthese* 189.1 (2012), pp. 97–118.

[121] Gideon Rosen. "Metaphysical Dependence: Grounding and Reduction". In: *Modality: Metaphysics, Logic, and Epistemology.* Ed. by Bob Hale and Aviv Hoffmann. Oxford University Press, 2010, pp. 109–36.

[122] Jonathan Schaffer. "Grounding, Transitivity, and Contrastivity". In: *Metaphysical Grounding: Understanding the Structure of Reality.* Ed. by Fabrice Correia and Benjamin Schnieder. Cambridge University Press, 2012, pp. 122–138.

[123] Jonathan Schaffer. "On What Grounds What". In: *Metametaphysics: New Essays on the Foundations of Ontology.* Ed. by David Manley, David J. Chalmers, and Ryan Wasserman. Oxford University Press, 2009, pp. 347–83.

[124] Benjamin Schnieder. "A Logic for 'Because'". In: *The Review of Symbolic Logic* 4.3 (2011), pp. 445–65.

[125] Benjamin Schnieder. "Canonical Property Designators". In: *American Philosophical Quarterly* 43.2 (2006), pp. 119–32.

[126] Peter Schröder-Heister. "A Natural Extension of Natural Deduction". In: *Journal of Symbolic Logic* 49.4 (1984), pp. 1284–300.

[127] Stewart Shapiro. *Foundations Without Foundationalism: A Case for Second-Order Logic.* Oxford University Press, 1991.

[128] Alexander Skiles. "Against Grounding Necessitarianism". In: *Erkenntnis* 80.4 (2015), pp. 717–51.

[129] Roy Sorensen. “Nothingness”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2015. 2015.

[130] Chris Swoyer and Francesco Orilia. “Properties”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Fall 2014. 2014.

[131] Alfred Tarski. “Der Wahrheitsbegriff in den Formalisierten Sprachen”. In: *Studia Philosophica* 1 (1936), pp. 261–405.

[132] Alfred Tarski. “The Semantic Conception of Truth: And the Foundations of Semantics”. In: *Philosophy and Phenomenological Research* 4.3 (1944), pp. 341–76.

[133] Kelly Trogdon. “An Introduction to Grounding”. In: *Varieties of Dependence. Ontological Dependence, Grounding, Supervenience, Repsonse-Dependence*. Ed. by Miguel Holtje, Benjamin Schnieder, and Alex Steinberg. Philosophia Verlag, 2013, pp. 97–122.

[134] Linda Wetzel. “Is Socrates Essentially a Man?” In: *Philosophical Studies* 98.2 (2000), pp. 203–20.

[135] Nathan Wildman. “Modality, Sparsity, and Essence”. In: *Philosophical Quarterly* 63.253 (2013), pp. 760–82.

[136] Timothy Williamson. *Modal Logic as Metaphysics*. Oxford University Press, 2013.

[137] Stephen Yablo. “Identity, Essence, and Indiscernibility”. In: *Journal of Philosophy* 84.6 (1987), pp. 293–314.

[138] Stephen Yablo. “Paradox Without Self–Reference”. In: *Analysis* 53.4 (1993), pp. 251–52.

[139] Edward N. Zalta. “Essence and Modality”. In: *Mind* 115.459 (2006), pp. 659–93.

[140] Edward N. Zalta. *Intensional Logic and the Metaphysics of Intentionality*. MIT Press, 1988.