NUMBERS AND NECESSITY

Craig Warmke

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Philosophy.

Chapel Hill
2015

Approved by:
Robert M. Adams
Thomas Hofweber
Keith Simmons
L.A. Paul
William G. Lycan
I offer a new account of property parthood, and on its basis, novel semantic approaches to first-order logic, modal propositional logic, and quantified modal logic. I also use it in my metaphysical accounts of sets and the natural numbers. Along the way, I analyze property parthood in terms of the mental conjunction of ideas. What results: idealist accounts of necessity and number with certain explanatory advantages over standard accounts.
For Laura…
ACKNOWLEDGEMENTS

I have been in grad school for nine years and accrued debts to people far and wide. Rather than list the entire ledger, I’ll tell a few stories instead.

First. When I applied to PhD programs, I emailed three people about working with them: Thomas Hofweber at UNC, Laurie Paul at Arizona, and Stephen Yablo at MIT. Only Arizona accepted me, but Laurie had offers to teach at other schools, including UNC. She very graciously promised to try to take me with her if she decided to go elsewhere. Soon after I agreed to attend Arizona, Laurie took the UNC job. But she stuck around Tucson for my first year, worked closely with me, and then, as promised, took me to UNC.

I arrived in Chapel Hill at the same time as Bob Adams. He taught a seminar on existence our first semester at UNC, and I wrote a paper for him which developed a new semantics for modal logic. Bob encouraged me to continue this project, which grew into the present dissertation. He co-supervised my dissertation for two years with Keith Simmons until they both left for the northeast. At that point, UNC rules required me to adopt another supervisor from UNC. Since I was working on logic and philosophy of mathematics at the time, Thomas Hofweber was the natural choice to step in for Keith.

For all this moving and shuffling, I’ve had the enviable experience of having four world-class advisers. And I owe them each a unique debt. I’m grateful for the lessons Laurie taught me at both Arizona and UNC and for letting me tag along. I’m grateful for Bob’s encouragement and depth of philosophical and historical insight. For Keith’s patience, availability, philosophical clarity, and formal precision. And for Thomas’s approachability and philosophical and professional guidance. UNC has been so important for my philosophical development. So I sincerely thank Stephen Yablo for ignoring my email and rejecting me from MIT.

Second. My brother Brandon and I went to the same small school in Chicago, but neither of us studied philosophy. After Brandon graduated, he began to claw his way into academic philosophy. He took some classes at some community colleges and eventually took a class with Daniel Sutherland...
at UIC. Partly on the basis of Sutherland’s letter of recommendation, my brother gained admission into the MA program at Northern Illinois. Probably on the basis of Brandon’s performance, I also was admitted to NIU. This fall, I’ll return to NIU as an Assistant Professor.

There are three people I’d like to thank here. First, my brother. He is my best friend and also an excellent philosopher. I wouldn’t be in philosophy without him. Being in the same profession as my brother has been one of my life’s greatest joys. Second, Tomis Kapitan. He taught me modal logic and possible worlds semantics at NIU. He retired recently, and, as I understand it, I’ll be filing the shoes he left behind—a fulfilling and challenging task. Finally, Daniel Sutherland. I might not be in philosophy without the letter he wrote for my brother. Amusingly, Sutherland and I are now also brothers of a sort, since Bob Adams has helped advise both of our dissertations.

Third. Laura and I married soon after I graduated from college. We spent our savings on gas and tuition so I could commute over an hour to attend graduate school in a field in which I had no background. She gave up job opportunities and moved cross country twice so I could study where I felt I should. She agreed to have two children with me in graduate school even though job prospects for philosophers are grim. Besides my parents, no one has believed in me more or loved me as much in such a costly way.

Like any dissertation, mine has a table of contents which captures the topics to come. But each entry in the TOC carries a special meaning for me beyond its normal one. I see a section and remember the life I had while I wrote it. I see ups and downs as I scan my TOC, and Laura’s love and support is a common thread through them all. So it’s only fitting for me to dedicate this whole thing to her.
# TABLE OF CONTENTS

LIST OF FIGURES ................................................................. x

1 Introduction ................................................................. 1

2 Property Parthood .......................................................... 7
   2.1 Introduction ......................................................... 7
   2.2 Abundant Properties ............................................. 9
   2.3 Candidate Analyses ............................................... 10
   2.4 Competing Conceptions .......................................... 14
   2.5 The Mereology ..................................................... 20
   2.6 Conclusion ........................................................ 25

3 Modal Intensionalism .................................................... 27
   3.1 Introduction ......................................................... 27
   3.2 Inclusion ............................................................ 29
   3.3 Frames ............................................................... 30
   3.4 Truth, Necessity, and Possibility .............................. 32
   3.5 Pure Semantics .................................................... 37
   3.6 Applied Semantics ............................................... 39
   3.7 Explanation and Finite Cognition .............................. 41
   3.8 Conclusion ........................................................ 44

4 Logical Intensionalism .................................................. 45
   4.1 Introduction ......................................................... 45
   4.2 Montague Semantics .............................................. 49
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>The Containment Theory</td>
<td>53</td>
</tr>
<tr>
<td>4.4</td>
<td>The Alternative</td>
<td>61</td>
</tr>
<tr>
<td>4.5</td>
<td>Evaluation</td>
<td>69</td>
</tr>
<tr>
<td>4.6</td>
<td>Conclusion</td>
<td>77</td>
</tr>
<tr>
<td>4.7</td>
<td>Appendix 1: Examples</td>
<td>78</td>
</tr>
<tr>
<td>4.8</td>
<td>Appendix 2: Non-existents</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>Modal Idealism</td>
<td>83</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>83</td>
</tr>
<tr>
<td>5.2</td>
<td>Ideas</td>
<td>87</td>
</tr>
<tr>
<td>5.3</td>
<td>Parts of Ideas</td>
<td>93</td>
</tr>
<tr>
<td>5.4</td>
<td>Worldhood</td>
<td>97</td>
</tr>
<tr>
<td>5.5</td>
<td>Propositional Ideas</td>
<td>100</td>
</tr>
<tr>
<td>5.6</td>
<td>Truth</td>
<td>103</td>
</tr>
<tr>
<td>5.7</td>
<td>Necessity</td>
<td>105</td>
</tr>
<tr>
<td>5.8</td>
<td>Names</td>
<td>106</td>
</tr>
<tr>
<td>5.9</td>
<td>Predicates</td>
<td>108</td>
</tr>
<tr>
<td>5.10</td>
<td>Negation</td>
<td>109</td>
</tr>
<tr>
<td>5.11</td>
<td>Quantification</td>
<td>112</td>
</tr>
<tr>
<td>5.12</td>
<td>The Formalism</td>
<td>113</td>
</tr>
<tr>
<td>5.13</td>
<td>Parthood</td>
<td>118</td>
</tr>
<tr>
<td>5.14</td>
<td>Malebranche</td>
<td>121</td>
</tr>
<tr>
<td>5.15</td>
<td>Divine Idealism</td>
<td>126</td>
</tr>
<tr>
<td>6</td>
<td>Set-theoretical Idealism</td>
<td>132</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>132</td>
</tr>
<tr>
<td>6.2</td>
<td>Preliminaries</td>
<td>133</td>
</tr>
<tr>
<td>6.3</td>
<td>Membership</td>
<td>135</td>
</tr>
<tr>
<td>6.4</td>
<td>Desiderata</td>
<td>138</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

3.1 Modal intensionalism (left) and modal extensionalism (right) ......................... 32
3.2 An intensionalist approach to truth. ....................................................... 34
3.3 Necessary truth according to modal extensionalism. ................................. 35
3.4 Necessary truth according to modal intensionalism. ................................. 35
5.1 The idea of a’s being F is part of the idea of the actual world. .................... 103
5.2 The idea of a’s being F is part of the idea of being a world. ....................... 105
CHAPTER 1

Introduction

Over the last 21 months, my daughter’s language and reasoning skills have exploded. She knows that dogs bark and birds chirp; that a single bean is *one*, a pair of beans is *two*, and a few beans are *three*; and, on some occasions, that she would have been in time out had she not behaved as well. She thinks about individuals and kinds and even numbers and possibilities.

“How does she do that?!”—when I ask this question as her father, I think of the difference between now and her first days outside the womb as barely conscious human pudding. When I ask this question as a philosopher, however, I wonder how my child’s cognitive periscope has broken through our concrete underworld of dogs and birds into the abstract realm of number and possibility. No one has laid eyes or fingers on numbers or encountered anything but actual objects. Yet somehow my child grasps both mathematical objects and modal truths about what’s possible or necessary. The world would be mysterious enough if she alone could grasp these things. But you and I do, too. How do we do it?

We can make progress on these questions if we turn back the clock a few hundred years. At a fork in the road long ago, philosophers took a turn towards extensional approaches to logic, meaning, and metaphysics. In contemporary logic, for example, the meaning of a predicate is its extension, something like the set or class of individuals which satisfy the predicate. And in linguistics, philosophy of language, and metaphysics, even purportedly intensional objects are treated as if they were large sets or classes. A property is the set of possible individuals which exemplify that property in any possible world; a concept is the set of possible individuals which satisfy the concept in any possible world. This extensional turn has helped in various ways, especially in calculating truth-values. But extensional approaches only provide functional descriptions of a deeper reality. They do not illuminate the metaphysics or epistemology of meanings or intensional entities generally.
Consider the predicate ‘is blue’, for example. What does it mean? Suppose a predicate’s meaning is the property it denotes. Then the predicate ‘is blue’ means the property of being blue. But the property of being blue isn’t a set. If properties were sets of actual individuals, then if all the blue things coincided perfectly with all the round things, the property of being blue would be identical to the property of being round. And properties are not sets of possible individuals either. Something is blue partly because it appears a certain way in certain conditions. Although we may define a set so that each member appears a certain way in certain conditions, being a member of a set doesn’t make anything appear as such-and-such in any conditions.

Or suppose a predicate’s meaning is the idea (or concept) of the property denoted by the predicate. But ideas are not sets of actual individuals. Otherwise, if all the blue things coincided perfectly with all the round things, the idea of being blue would be the idea of being round, which is absurd. If the idea of being blue were a set, it would be the set of all possible blue individuals. But when we think of being blue, we do not, and can not, think of so many individual blue objects. We have a fairly good grasp of the idea of the property of being blue, but the experience of grasping this idea does not at all seem like grasping a set of possible individuals. Clever philosophers can massage these problems away, at least initially. That anyone would want to, however, shows just how far down we’ve gone down the extensional road.

Let’s double back to the fork in the road and see where we might have gone instead. Back at the fork, Leibniz describes two inverse approaches to logic:

For when I say *Every man is an animal* I mean that all the men are included amongst the animals; but at the same time I mean that the idea of animal is included in the idea of man. ‘Animal’ comprises more individuals than ‘man’ does, but ‘man’ comprises more ideas or more attributes: one has more instances, the other more degrees of reality; one has the greater extension, the other the greater intension.¹

When Leibniz considers the proposition that every man is an animal, he has in mind all possible men. When we treat such a proposition extensionally, we’ll say that it is true just in case the set of possible men is included (as a subset) in the set of possible animals. But we could also treat the proposition intensionally and say that it’s true just in case the idea of being a man includes the idea of being

¹NE, IV, xvii, 8, p. 486.
an animal. Leibniz correctly notes in the passage that the intensional and extensional approaches are inversely related. Extensionally, ‘is an animal’ means the set of all animals, which includes (as subsets) all sets of each different kind of animal, including the set of men or human beings. Intensionally, ‘is a man’ means the idea of being a man, which include ideas of the preconditions for being a human in general, including the ideas of being an animal and being a mammal.

Leibniz offers us a more psychologically realistic account of the meanings of predicates. When I grasp the meaning of ‘is a human’, what I grasp is either the property of being human itself or the idea of being human. And in grasping the property or the idea, I do not grasp a set of actual or possible individuals, but more properties or ideas: the property or idea of being a mammal, the property or idea of being an animal, and so on. That is how it seems to me anyway. I have good evidence that it seems the same to you, too. The Leibnizian structure of thought and meaning overflows into our discourse. We say things like “being a mammal is part of being human” and “being red is part of being crimson.” In fact, googling “part of being” provides scores of similar examples (though you’ll have to wade through a non-negligible amount of material about Heidegger). A preponderance of such judgements is exactly what one would expect if Leibniz was more or less right about the structure of intensional entities. I propose we take these judgments at face-value and provide a formal theory of property parthood in Chapter 1. On the basis of that theory, we can develop intensional approaches to logic, modality, and mathematics.

Since the intensional and extensional approaches are inversely related, one should suspect that an intensional approach for some domain could accomplish at least as much as its standard extensional approach. To some extent, my dissertation bears out this suspicion. But I don’t primarily aim to prove that intensional approaches are formally adequate. I primarily aim to show that intensional approaches are better in some ways than standard extensional approaches. Even though intensional and extensional approaches are, in some ways, formally isomorphic, intensional approaches enjoy advantages due to their different formal structures. These different formal structures are ripe for certain metaphysical interpretations, and the pairing of these structures with one such interpretation provides powerful explanations of what modal reality and mathematical reality consist in and how we know about them.

The structure of intensional approaches helps explain how we make judgments in the ways that we do. Extentional approaches have divorced the ways we make judgments from their meaning.
An intensional approach can help reunite judgment to meaning. Take the simple case of singular propositions, for example. Fred is a guy and being tall is a property. So when we judge that Fred is tall, we judge that some guy has some property. And when we judge that Fred is tall, our idea of Fred includes being tall, an idea of the property of being tall. So we have two candidates for the meaning of the predicate ‘is tall’. On the one hand, we have the property of being tall. On the other hand, we have the idea of that property. Yet in our first logic classes, we learn that ‘is tall’ means the collection of all things tall and that ‘Fred is tall’ is true just in case the collection contains Fred. The standard approach assigns the wrong kind of object for the meanings of predicates. And, as a result, the standard approach uses the wrong direction of containment. Properties and ideas are intensional entities and we will want another semantics for first-order logic if intensional entities are the meanings of predicates. In Chapter 3, I offer a semantics for first-order logic according to which ‘Fred is tall’ is true just in case an intensional entity, being Fred, other includes an intensional entity, being tall.

Or consider the proposition that all mammals are animals, for instance. How do we so confidently judge that the proposition is true without surveying every single mammal on the planet? From our observations with many mammals, we build the idea of being an animal into the idea of being mammalian. If this idea of being mammalian accurately represents what it is to be a mammal, then nothing could satisfy the idea of being a mammal without also being an animal. We can correctly judge that all mammals are animals without surveying all mammals once we inspect the idea of being mammal and find the idea of being an animal within it.

Or consider some proposition \( p \) which we judge to be necessarily true. Standardly, \( p \) is necessarily true just in case it is true in all possible worlds. But presumably there are infinitely many possible worlds. How could we judge that something is true in every one of these infinitely many possible worlds? First, rather than define necessity as truth in possible worlds, we can define necessary truths as preconditions for the existence of a world in general: \( p \) is necessary when being such that \( p \) is part of being a world in general. (I develop an entirely new modal semantics around this definition of necessity in Chapters 2 and 4.) The new definition of necessity explains how we can make judgments about all possible worlds. Any possible world, if it had been actual, would have exemplified the property (or satisfied the idea) of being a world in general. So if being such that \( p \) is part of being a world in general, then any possible world, if it had been actual, would have exemplified the property
(or satisfied the idea) of *being such that* $p$. The propositions which are true in all possible worlds correspond to propositional properties or ideas which any possible world would have exemplified or satisfied if it had been actual. We can judge that $p$ is necessary or true in all possible worlds because we can judge that *being such that* $p$ is part of being a world in general.

For reasons I shall explain shortly, I favor ideas over mind-independent properties in my overall semantic machinery. The structure of ideas also helps to explain our grasping of sets and our ability to count small collections. Each idea of an individual includes an idea of being something (in general) and an idea of being the particular thing it is. We can form collective ideas of individuals, e.g., the idea of Tom, Dick and Harry. The set of Tom, Dick, and Harry, is the collective idea of the particular thing Tom, the particular thing Dick, and the particular thing Harry. In Chapter 5, I argue for the view that the empty set is the idea of nothing, and, in general, sets consist in a specific type of collective idea of individuals. Idealism about sets dissolves a number of puzzles about sets and has enough explanatory power to be taken seriously. We can abstract from the collective idea of Tom, Dick, and Harry to the idea of something, something else, and another something, the number 3. In Chapter 6, I defend this idealist twist on the traditional units view of number against Fregean criticisms. I believe numerical idealism best captures the phenomenology of counting.

So why do I favor ideas instead of mind-independent properties? Many of the accounts here could go through with properties rather than ideas. But when we theorize about modality and mathematics, it seems to me that we inspect relationships among ideas. Placing modal and mathematical reality within mental space partially resolves Benacerraf-style challenges to our access to modal and mathematical truths. If modal and mathematical truths consist in relationships between ideas, which are mind-dependent, it is not so puzzling how we have access them. In Chapter 4, I defend a kind of Malebrancheanism about ideas according to which ideas are causally efficacious extensions of an infinite and omnipotent mind. If mathematical objects and modal truths consist in causally efficacious ideas, we can appeal to causation to explain our connection to modal and mathematical reality. Chapters 5 and 6 on sets and numbers are written with the same view in mind.

My dissertation is an exploration of idealism about modality and mathematics. It is not concerned with idealism about material objects. Idealist accounts, and intensional accounts generally, have some explanatory advantages over standard accounts. I probably haven’t done them justice. So I hope the reader can appreciate my own assessment of what follows. The key passages on idealism occur at the
beginning and end of Chapter 4. But I’m quite dissatisfied with the formal presentation in Chapter 4 and the structure and clarity of it and Chapter 6. Chapter 2 is my best and most important work. The other chapters have redeemable qualities, especially Chapter 3. I hope the chapters here inspire someone to question not only their commitments but also the value of vast swaths of contemporary metaphysics. That is the best I can hope for as a metaphysician.
CHAPTER 2

Property Parthood

2.1 Introduction

We say of many different types of thing that they have parts.¹ In the familiar spatio-temporal sense of ‘part’, desks and cups and cats and dogs have parts.² Or at least we talk as if they do. But there are senses of ‘part’ besides the spatio-temporal one.³ We also say that being an animal is part of being a mammal, that being red is part of being crimson, and that being four-sided is part of being square.⁴ This way of speaking is both common and meaningful. If you’re skeptical, search “part of being” in Google and you’ll see scores of similar judgments. (Beware of the considerable amount of unrelated results about Heidegger.) Philosophers speak this way, too. I could provide a long appendix of quotations, but two quotations will suffice:

Part of being a cocktail party is being thought to be a cocktail party; part of being a war is being thought to be a war.⁵

On this view, being charitable is part of being fair to those with whom one interacts.⁶

Being crimson and being square are what we would ordinarily call “properties,” and our property parthood judgments differ in important ways from other ordinary parthood judgments. We say things like “being red is part of being crimson” and “being a mammal is part of being a dog.” But redness is

¹See Winston et al. (1987), Iris et al. (1988), Gerstl and Primmenow (1995), and especially Varzi (2009) for examples.
²By this or that sense of ‘part’, I mean no more than the way we say that this or that type of thing has parts. Different senses of ‘part’ may point to different kinds of parthood, but I’m neutral on that issue here.
³Lewis (1991, 75–76) and Armstrong (1978, 37 ff.). Rosenkrantz and Hoffman (1991, 843) defend a generic conception of ‘parthood’ which has spatial, temporal, and logical parthood as its species.
⁴Both Varzi (2009) and Armstrong (1978, 37–38) provide similar examples.
exemplified by every crimson thing and by objects in many other places besides. Similarly, mammals are located everywhere dogs are and in more places besides. In these cases, the putative whole occupies less space-time than the putative part—if properties occupy space-time at all. But objects not only seem to occupy space-time, they also seem to occupy at least as much space-time as any one of their parts. It would be surprising to find a mug which occupied less space-time than its handle, for example.

The kinds of property parthood judgments under consideration may betray nothing more than a loose and popular way of speaking. But it may be worth taking more seriously. Like many judgments about objects and their parts, many property parthood judgments seem true. But what in the world would make them true? Perhaps property parthood judgments somewhat reliably track some other relation among properties. If so, what would that other relationship be? My attempts to uncover such a relation, which I shall explain later, have convinced me that property parthood judgments somewhat reliably track an irreducible property parthood relation. Positing such a relation raises a host of other questions. What principles would govern such a relation? How would we have epistemic access to whether one property is part of another? What, if any, are the theoretical payoffs of positing such a relation?

In this first chapter, I carve out a conception of property parthood, argue that it is distinct from other available conceptions of property parthood, defend it against the worry that property parthood is incoherent, offer an axiomatic theory for the parthood relation.7 Classical Extensional Mereology (CEM), the formal system that many use to model spatio-temporal parthood, is widely used and studied, so I will use CEM as both marker and measuring stick for the property mereology. In some respects, the property mereology closely resembles CEM. But they diverge from each other in interesting ways. Although the property mereology is intrinsically interesting in its own right, it also plays an instrumental role in the following chapters. On its basis, I construct a novel semantics for first-order logic, another for modal propositional logic, and another for quantified modal logic.

7My property mereology is similar in spirit to the mereology of L. A. Paul (2002, 2006). I compare and contrast Paul’s property mereology with my own in section X.
2.2 Abundant Properties

When I say that properties have parts, I have in mind an abundant and not a sparse conception of properties. According to Lewis (1986b, 60), sparse properties account for

...qualitative similarity, they carve at the joints, they are intrinsic, they are highly specific, the sets of their instances are ipso facto not entirely miscellaneous, there are only just enough of them to characterise things completely and without redundancy.

On the abundant conception, however, there is a property corresponding to every predicate.\(^8\) The predicate ‘is red or blue’ expresses the disjunctive property being red or blue and ‘is five miles from New York’ expresses the extrinsic property being five miles from New York. When two things share the disjunctive property of being red or blue, for instance, they may do so because one is red and the other blue. Sharing an abundant property doesn’t guarantee qualitative similarity.

True to its name, the abundant conception licenses more than just disjunctive and extrinsic properties. We say things like “three is not even” and “three is odd and prime” using the predicates ‘is not even’ and ‘is odd and prime’. So on the abundant conception, there are negative properties like being not-even and conjunctive properties like being odd and prime. We also sometimes refer to inconsistent properties that nothing could exemplify, like being round and square or being a set of all sets that are not members of themselves. The abundant conception licenses their existence, too.\(^9\)

Besides disjunctive, conjunctive, negative, and inconsistent properties, the abundant conception also countenances propositional properties. We can also prefix ‘is such that’ to any sentence to form a predicate that expresses a propositional property. Prefix ‘is such that’ to ‘three is odd’ and the result, ‘is such that three is odd’, expresses the propositional property being such that three is odd.\(^10\) There

---

\(^8\) Lewis (1986b, 59–60)

\(^9\) By similar reasoning, one may conclude that there is a property of being non-self-exemplified. But does it exemplify itself? If it does, then it would seem to have the property of non-self- exemplification and so not exemplify itself. But if it does not exemplify itself, it would seem to have the property of non-self-exemplification and so exemplify itself. van Inwagen (2004, 135–136) observes a similar problem with his commitment to properties as assertibles. van Inwagen prefers to deny that there is any such problematic property but doesn’t endorse any particular strategy for doing so. I would prefer to accept the problematic property, but I have no strategy for avoiding paradox. Thanks to Joshua Rasmussen for discussion on this point.

\(^10\) In Chapter 3, Section 4, I discuss the differences between propositional properties and non-propositional properties in more detail.
are mixtures, too: disjunctive properties whose disjuncts are propositional properties, conjunctive properties comprised of negative properties, and so on.

Abundant properties abound, to be sure. Those with predilections for desert landscapes may wonder whether the abundant conception countenances too many properties. I accept abundant properties because they play such a useful role in intensional semantic approaches to logic and modality, which are preferable in some ways to their extensionalist counterparts.\textsuperscript{11} For me, then, whether the price of abundant properties is worth their plunder hinges on, first, how metaphysically controversial they are, and, second, how successful intensional approaches to logic and modality are.

Let’s visit that first point, how controversial these abundant properties are. Insofar as I can, I’ll remain neutral about the metaphysics of abundant properties. The point of the present chapter isn’t to provide a specific metaphysics of the properties which appear in our property parthood judgments. Rather, the point is to develop a loosely interpreted formal theory of property parthood which explicitly captures our the kinds of property parthood judgments with which we began. The metaphysics can wait.\textsuperscript{12}

Yet I cannot be fully neutral about the metaphysics of abundant properties. Their abundance alone precludes what I’ll call \textit{Aristotelianism}, the view that properties are object-dependent and do not exist unless they are exemplified by an object. Aristotelianism implies that there are no unexemplified properties. If nothing had ever exemplified redness, let’s say, it wouldn’t have existed. \textit{Platonism} denies that properties are object-dependent and thus allows for unexemplified properties. I endorse Platonism about abundant properties and don’t mind if many properties never punch their tickets as properties \textit{of} something. Aristotelianism may be consistent with some form of modal intensionalism, but the approach is more easily expressed with Platonism’s extra resources.

\subsection*{2.3 Candidate Analyses}

When I say that properties have \textit{parts}, I have in mind a conception of parthood that would capture many ordinary parthood judgments.\textsuperscript{13} We judge that being mammalian is part of being a dog, for

\textsuperscript{11}I develop intensional approaches to first-order logic and modal logic in Chapters 2–4. Lewis (1983, 16–19) argues that abundant properties are necessary for sketching a systematic semantics and for characterizing our intentional attitudes.

\textsuperscript{12}Until Chapter 4, that is.

\textsuperscript{13}This section and the next borrow heavily from my paper “Property Parthood.”
example. The most obvious candidate analyses of such judgments associate the part with some sort of necessary condition of anything’s exemplifying the whole. Here I survey a handful of these analyses and explain why they are less than plausible.

The *necessary condition* analysis says that *being F* is part of *being G* if and only if whatever is G is also F. On this analysis, saying that a property is part of another is shorthand for saying that exemplifying the former is a necessary condition of exemplifying the latter. The necessary condition analysis captures an important aspect of some parthood judgments, but it has an implausible consequence. The important aspect of some parthood judgments is their asymmetry: ordinarily, we judge that being mammalian is part of being a dog but not that being a dog is part of being mammalian. The necessary condition analysis nicely captures this asymmetry. Whatever is a dog is also a mammal, so with the schema we may infer that being mammalian is part of being a dog. But not all mammals are dogs, so the schema doesn’t allow us to infer that being a dog is part of being a mammal.

The implausible consequence of the necessary condition analysis concerns unexemplified properties. Consider an unexemplified property such as the property of being a gold mountain. True: every golden mountain is also not-gold. Why: there are no gold mountains, so the conditional is vacuously true. The vacuously true conditional and the schema wrongly imply that *being not-gold* is part of *being a golden mountain*. Although there are no golden mountains, there could be. And there could be golden mountains only if the property of being a golden mountain doesn’t have parts whose joint exemplification would be impossible. The necessary condition analysis would allow us to infer that every unexemplified property were necessarily unexemplified in a similar way. While many unexemplified properties are necessarily unexemplified, an analysis of property parthood which implies that every unexemplified property is necessarily unexemplified is too steep a price to pay.

Instead of associating a property’s parts with necessary conditions for its exemplification, perhaps we should associate them with necessarily necessary conditions in a modal analysis of property parthood. On such an analysis, *being F* is part of *being G* if and only if, necessarily, every G is also F. This schema captures many of our ordinary judgements about property parthood. First, necessarily, whatever is a dog is also a mammal, so the modal analysis correctly implies that being mammalian is part of being a dog. Second, it isn’t the case that, necessarily, whatever is a mammal is also a dog. So the modal analysis does not imply that being a dog is part of being mammalian. Like the necessary
condition analysis, the modal analysis captures this important asymmetry these kinds of property parthood judgments.

And unlike the necessary condition analysis, the modal analysis fails to imply that unexemplified properties are necessarily unexemplified. Presumably, it’s not true that, necessarily, every golden mountain is not-gold. Every possible golden mountain is gold, not not-gold. The schema doesn’t imply that being non-gold is part of being a gold mountain. Since the modal schema doesn’t imply that being (not-)G is part of being F given a vacuously true conditional that every F is (not-)G, it doesn’t imply that unexemplified properties are necessarily exemplified.

Despite its advantages over the necessary condition analysis, the modal analysis implies a number of controversial theses about properties. Traditionally, \(x\) is part of \(y\) just in case \(x\) is either identical to \(y\) or \(x\) is a proper part of \(y\). Given that something cannot be a proper part of one of its own proper parts, saying that \(x\) and \(y\) are parts of each other is a roundabout way of saying that \(x\) is identical to \(y\), given the traditional characterization of parthood. And given that mutual parthood implies identity, the modal analysis implies that necessarily coextensive properties are identical. For, necessarily, something is trilateral if and only if it is triangular. Therefore, the modal analysis implies that being trilateral and being triangular are mutual parts and therefore identical. We can reason similarly to the identity of the propositional properties being such that no contradiction is true and being such that \(2 + 2 = 4\). Identity claims about properties such as these are not obvious, and to the minds of many they are dubious.

The modal analysis also implies controversial theses about necessarily unexemplified properties. Consider the property of being red and not-red. Necessarily, there is no thing that is both red and not-red. So, in terms of possible worlds: in each possible world, every instance of the schema ‘whatever is red and not-red is F ’ is vacuously true. As a result, every instance of the schema ‘necessarily, whatever is red and not-red is F’ is true. So the modal analysis implies that every property whatever is part of being red and not-red. And the same is true for every necessarily exemplified property. If necessarily unexemplified properties have all properties as parts, then they have all the same parts. Traditionally, when \(x\) and \(y\) share all the same parts, \(x\) and \(y\) are identical. Therefore, given the traditional view that complete overlap of parts makes for identity, the schema implies that there is only one necessarily unexemplified property: being a round square, being such that \(2 + 2 = 3\), and
being red all over and green all over are the same.\textsuperscript{14} That there is only one necessarily unexemplified property is not obvious. And if we want properties to serve as the semantic values of predicates, including predicates which seem to express distinct, necessarily unexemplified properties, we will want many different necessarily unexemplified properties.

Perhaps a parthood judgment of the form being \( F \) is part of being \( G \) is little more than a judgment that whatever is \( G \) is thereby also \( F \)—that whatever is \( G \) is, in virtue of being \( G \), also \( F \). But ‘in virtue of’ is notoriously opaque. Do we gain any understanding of property parthood by analyzing property parthood judgments with an opaque instance ‘in virtue of’? Saying “whatever is \( G \) is, in virtue of being \( G \), also \( F \)” is, first, more complicated and than saying that being \( F \) part of being \( G \) since the former concerns the relation between an individual’s exemplifying a property and an individual’s exemplifying a property, whereas the property parthood judgment concerns no individuals or exemplification relations, but only the relation between two properties. The ‘in virtue of’ judgment is not only more complicated, but unnecessarily so. What makes it true that whatever is \( G \) is, in virtue of being \( G \), also \( F \), if not some relation between the properties being \( F \) and being \( G \)? If being \( F \) is part of being \( G \), then anything’s exemplifying the whole involves its exemplifying the whole’s parts. Whatever is \( G \) is, in virtue of being \( G \), also \( F \) precisely because being \( F \) is part of being \( G \).

When we say that being \( F \) is part of being \( G \) we might mean that being \( F \) is “part of what it takes” to be \( G \).\textsuperscript{15} Being red is part of what it takes to be crimson, but the intended sense here doesn’t mean that being red is merely a necessary condition for anything’s being crimson. Something’s being \( F \) is a necessary condition of its being \( G \) when either (i) anything is \( G \) only if it is also \( F \) or when (ii) necessarily, anything is \( G \) only if it is also \( F \). Analyzing property parthood via (i) leads to the incorrect non-modal schema for property parthood, and analyzing it via (ii) leads directly to the modal schema for property parthood. I’ve rejected both of these analyses already. The “part of what it takes” construal of property parthood is likely helpful only insofar as we lacked a feel for the conception(s) of property parthood guiding many of our judgments about properties.

\textsuperscript{14}In Chapter 2, I explicitly reject the modal analysis because I use property parthood in a semantics for first-order logic. The semantics for first-order logic says that ‘Fred is tall’ is true when being tall is part of being Fred. But Fred isn’t thereby necessarily tall. In Chapters 3 and 4, I define modal notions in terms of property parthood rather than the other way around. So the modal notions some might use to define property parthood I define in terms of parthood relations between properties.

\textsuperscript{15}Recently, Rosen (2010) has used the “part of what it takes” locution while discussing the grounding relation.
The desire for an analysis of property parthood is understandable, and I offer one in Chapter 4. But an analysis isn’t necessary to motivate and develop the property mereology itself. Our intuitive judgments about property parthood motivate the mereology, and our intuitive grasp of those judgments is enough for mapping out some of its formal features. Instead of an analysis, I here offer an axiomatic theory as partly constitutive of a conception of property parthood which captures many of our property parthood judgments.

### 2.4 Competing Conceptions

The most well-known philosophical conception of property parthood occurs in discussions about structural universals. Structural universals are universals because they are wholly located wherever they are exemplified. They are structural because they consist of more than one universal. The structure of such a universal includes whatever universals a thing’s spatio-temporal proper parts exemplify in virtue of exemplifying that universal. On the structural conception of property parthood, the parts of a universal are those universals a thing’s spatio-temporal proper parts exemplify in virtue of exemplifying that universal.\(^\text{16}\) An H\(_2\)O molecule is composed of two hydrogen atoms and an oxygen atom. When something exemplifies the universal of being an H\(_2\)O molecule, some of its spatio-temporal proper parts exemplify the universals being hydrogen and being oxygen. Given the structural conception of parthood, they are parts of being an H\(_2\)O molecule.

Some universals may appear multiple times in the structure of another, as being hydrogen does in being H\(_2\)O, for example. And some pairs of universals differ not in which universals appear in their structures but in the number of times those universals appear in their respective structures. For instance, being carbon and being hydrogen both appear in the structures of being methane and being butane. But being carbon appears once in being methane and four times in being butane; being hydrogen appears ten times in being butane and four times in being methane. Consequently, the controversy surrounding the structural conception of parthood concerns whether something can be

\(^{16}\) Lewis (1986a, 33).
a part multiple times over\textsuperscript{17} and whether different universals could be composed of the very same parts.\textsuperscript{18}

Lewis (1986a) worries that the structural conception of property parthood is incoherent. If the structural conception is incoherent, as Lewis worries, is my preferred conception of property parthood incoherent, too? Here is an instance of the worry: is being oxygen part of being an $H_2O_2$ molecule once or twice? On the one hand, being oxygen is a universal that is wholly located in all its instances, so the answer seems to be “once.” But this reasoning threatens to erase the structural differences of molecules which share the same kinds but not the same number of atoms, like $H_2O$ and $H_2O_2$. For if being oxygen is part of being $H_2O_2$ exactly once, what differentiates being $H_2O_2$ from being $H_2O$, which has being oxygen as a part exactly once? On the other hand, being oxygen appears twice in being an $H_2O_2$ molecule’s structure, so it seems to be a part of being $H_2O_2$ twice over. But what are there two of, since being oxygen is wholly present in each of its instances?

My conception of property parthood evades Lewis’s worry. First, unlike structural universals, the abundant properties which I have in mind are not present in their instances, let alone wholly present. Second, my conception of parthood disagrees with the structural conception about which properties are parts of which. Structural parthood is neither necessary nor sufficient for parthood in my sense. Consider the necessity claim, first. Being an atom is part of being an oxygen atom on my conception. Being an atom is part of what it takes to be an oxygen atom, we might say. But being an atom is not part of being an oxygen atom on the structural conception. If it were, then, by the definition of parthood, being an atom would either be identical to or a proper part of being an oxygen atom. Being an atom isn’t identical to being an oxygen atom since there are non-oxygen atoms. Being an atom isn’t a proper part of being an oxygen atom on the structural conception either. If it were, then, by that very conception, some spatio-temporal proper part of an oxygen atom would exemplify being an atom in virtue of exemplifying the universal of being an oxygen atom. Which spatio-temporal proper part of an oxygen atom would that be? No atom has an atom as a spatiotemporal proper part. Since no spatio-temporal proper part of an oxygen atom exemplifies being an atom, being an atom is not part of being an oxygen atom on the structural conception.

\textsuperscript{17} Lewis (1986a) and Bennett (2013).
\textsuperscript{18} Hawley (2010).
Now consider the sufficiency claim. *Being oxygen* is part of *being an H\textsubscript{2}O molecule* on the structural conception, since whatever exemplifies *being an H\textsubscript{2}O molecule* has a spatio-temporal proper part which exemplifies *being oxygen*. But *being oxygen* isn’t part of *being an H\textsubscript{2}O molecule* in my preferred sense of ‘part’. If it were, then H\textsubscript{2}O molecules would be oxygen atoms. So there’s no question about whether *being hydrogen* is part of *being an H\textsubscript{2}O molecule* once or twice. However, *having a hydrogen atom as a spatio-temporal part* may very well be part of *being an H\textsubscript{2}O molecule* on my conception. Hydrogen atoms are spatio-temporal parts of H\textsubscript{2}O molecules. Since an H\textsubscript{2}O molecule has two hydrogen atoms as spatio-temporal parts, we may raise a worry similar to Lewis’s: is *having a hydrogen atom as a spatio-temporal part* part of *being an H\textsubscript{2}O molecule* once or twice?

Consider a specific H\textsubscript{2}O molecule, “Harry,” whose two hydrogen atoms are “Matthew” and “Carl.” *Having a hydrogen atom as a spatio-temporal part* is part of having any particular hydrogen atom as a spatio-temporal part. So *having a hydrogen atom as a spatio-temporal part* is part of both *having Matthew as a spatio-temporal part* and *having Carl as a spatio-temporal part*. So these latter two properties overlap. Furthermore, *having Matthew as a spatio-temporal part* and *having Carl as a spatio-temporal part* are both parts of *being Harry*. Given the transitivity of parthood, there are two inference chains to the conclusion that *having a hydrogen atom as a spatio-temporal part* is part of *being Harry*. Since *having Matthew as a spatio-temporal part* and *having Carl as a spatio-temporal part* overlap with respect to *having a hydrogen atom as a spatio-temporal part*, these two inference chains do not put any pressure on us to admit that the latter property is part of *being Harry* twice over. What is “twice” is not how many times *having a hydrogen atom as a spatio-temporal part* is part of *being Harry*, but rather how many properties of having this or that particular hydrogen atom as a spatio-temporal part are part of *being Harry* but also overlap with respect to *having a hydrogen atom as a spatio-temporal part*.\(^{19}\) The situation is similar with overlapping physical objects. If object \(x\) has \(y\) and \(z\) as proper parts, and \(y\) and \(z\) themselves overlap with respect to \(w\), then we may infer that

---

\(^{19}\)I suspect Lewis would agree with something like this response. Lewis (1986a, 41, n. 21) mentions briefly that there is room for a mereology of conjunctive universals, where the conjuncts are present wherever the conjunction is. He says that this picture is “quite natural” and that “he has no quarrel” with any of it. Though my properties are not universals, my preferred conception is amenable to a conjunctive analysis. In fact, I endorse a conjunctive analysis in Chapter 4. If you want, we can say that *being F* is part of *being G* just in case *being G* is a conjunction of properties, one of which is *being F*. My preferred conception owes much to Leibniz’s mereological view of concepts, and Leibniz (G vii 239/F 135) himself endorses the biconditional that the concept F is identical to the conjunction of F and G if and only if the conjunctive concept F contains G. Armstrong (1978, 39)’s argument against what he calls the Conjunction Principle (“if F is a property, and G is a property which is part of that property, then whatever particulars have F also have G”) merits a similar response.
\(w\) is part of \(x\) via its being part of \(y\) or via its being part of \(z\)—the two inference chains signify how many of \(x\)’s proper parts (besides \(w\)) overlap with respect to \(w\), not how many times \(w\) is part of \(x\).

We should also distinguish the structural conception from my preferred conception because the structural conception does not capture many of our ordinary parthood locutions concerning properties. Notice how unnatural it would be in many contexts to say that \textit{being chlorine} is part of \textit{being an HCl molecule}. In some contexts, we would hesitate to say this because we know that HCl molecules aren’t chlorine atoms. However, in those same contexts, we could comfortably say that \textit{having a chlorine atom as a spatio-temporal part} is part of \textit{being an HCl molecule}. If \textit{having a chlorine atom as a spatio-temporal part} is part of \textit{being an HCl molecule}, then anything with the property of being an HCl molecule also has the property of having a chlorine atom (as a spatio-temporal part). This seems right.

Given the intelligibility of my preferred sense of property parthood, it will be beneficial to compare and contrast other available conceptions of property parthood with my own. My conception is similar in spirit to L. A. Paul’s.\(^{20}\) We each construct a property mereology and then attempt to implement it in metaphysically interesting ways. The major differences are the kinds of properties we implement and how we implement them. Paul (2006, 634) endorses a “relatively sparse” Aristotelian conception of properties and uses her property mereology in a bundle theory of material objects. I endorse a Platonistic conception of abundant properties and later (in Chapters 2, 3, and 4) use my property mereology in semantic accounts of logic and modality. My mereology and Paul’s are formally similar, but our differences in approach lead to three further differences. First, due to the types of properties each mereology governs, I endorse unrestricted composition and Paul (2006, 634) does not. Second, my mereology contains two additional axioms that govern properties of properties. Finally, my mereology is meant to capture a way we often speak about properties; Paul’s is not. So whereas my account says that \textit{being red} is part of \textit{being crimson} and implies that determinables are parts of their determinates generally, Paul’s account is compatible with the view that determinates are parts of their determinables.

On the \textit{Boolean conception} of parthood, properties compose other properties through analogues of the Boolean operations of conjunction, disjunction, and negation. According to Dean Zimmerman

we can reach a property’s parts by “successive eliminations of disjuncts, conjuncts, and
the ontological analogue of the negation operator.” Rosenkrantz and Hoffman (1991, 845) endorse a
similar conception, according to which “a conjunctive property has each of its conjuncts as a logical
part, a disjunctive property has each of its disjuncts as a logical part, and so forth.”

The Boolean conception implies that being red is part of being not-red. This isn’t something
we’d ordinarily say, not in the same sense of ‘part’ we’d say that being a mammal is part of being a
dog. We wouldn’t ordinarily say that being red is part of being not-red because we know that red
things are not not-red. The Boolean conception also implies that being red is part of being red or
blue. But we wouldn’t ordinarily say this either because many things have the property of being red
or blue without being red. Finally, the Boolean conception says that logically simple properties (with
no conjuncts, disjuncts, etc.) like being crimson have no parts. On the present account, however,
logical simplicity has no bearing on whether a property has parts. Being crimson, though logically
simple, is not mereologically simple. Its parts may include properties like being red or having a
certain brightness or hue.

My focus is the conception of parthood that underlies many of our parthood judgments, including
the judgment that being red is part of being crimson. The final conception of property parthood, the
set-theoretic account of properties, has it the other way around, at least insofar as it can say that
properties are parts of properties. On the set-theoretic account, each property is a set of possible
individuals: being red is the set of red things across possible worlds, and being crimson is the set of
crimson things across possible worlds. Because whatever is crimson is red and not vice versa, the set
of possible crimson individuals is a subset of the set of possible red individuals. So if one set is part
of the other, the set of possible crimson things is part of the set of possible red things, not the other
way around.21

This is little more than Leibniz’s point—tidied up in contemporary terms—that there is an inverse
relation between a predicate’s intension and its extension.22 For Leibniz, a predicate’s extension
consists of the things to which the predicate truly applies. The extension of the predicate ‘is male’
then consists of all males. A predicate’s intension is the concept which everything in the extension
satisfies. Now consider the predicate ‘is male and unmarried’. Its extension is now smaller: the

---

21 Lewis (1991, 3–5) argues that the subclasses of a class are its only parts.
22 See C53 in Couturat (1903), P20 in Parkinson (1966), and especially Ch. 2 of Adams (1994).
unmarried males are among but also outnumbered by all males in general. Its intension is now larger: the concept of being male and unmarried contains not only the concept of being male but also the concept of being unmarried. So the intension of ‘is male and unmarried’ includes the intension of ‘is male’ but the extension of ‘is male’ includes the extension ‘is male and unmarried’.

Leibniz’s point applies equally well if we identify the intension of a predicate with a property composed of other properties and its extension with the set of individuals which satisfy the predicate across possible worlds. Then the intension of ‘is male’ is the property of being male, which is part of the property of being male and unmarried. The extension of ‘is male and unmarried’ is then the set of possible males who are unmarried, which is a subset of the set of possible males. So whereas the intension of ‘is male and unmarried’ includes the intension of ‘is male’, the extension of ‘is male’ includes (in the sense of having as a subset) the extension of ‘is male and unmarried’. The parthood relations we often attribute to properties maps fairly closely onto Leibniz’s intensional treatment of predicates. So the commitment to a mereology for properties in accordance with this way of speaking is little more than a commitment to an intensional approach to properties, in Leibniz’s sense. Such a commitment doesn’t preclude someone from adopting the existence of extensions of predicates across possible worlds. But it is an additional commitment to intensional entities with a certain kind of hierarchical formal structure.

Now it should also be clear why a commitment to the intensional approach also rules out the view that properties are functions from possible worlds to sets of individuals. Because a function is a set of ordered n-tuples, a monadic property on this view is a set of ordered pairs, where each ordered pair consists of a possible world, on the one hand, and the extension of the property in that world, on the other. The property of being red, for instance, is the set of all ordered pairs \( \langle w_i, S_i \rangle \) such that for each world \( w_i \), every member of \( S_i \) is red in \( w_i \). The function view doesn’t account for our common parthood attributions either. We commonly say things like, “being red is part of being crimson.” But, on this view, in order for being red to be part of being crimson, the set of ordered pairs which is the property of being red, \( R \), would have to be a subset of the set of ordered pairs which is the property of being crimson, \( C \). It isn’t, though: there are ordered pairs in \( R \) that are not members of \( C \). For in many possible worlds, the extension of crimson is a proper subset of the extension of red, so the ordered pairs these worlds and their extensions in \( R \) are not members of \( C \). Hence, a commitment
to Leibniz’s intensional approach applied to properties rules out both the set-theoretic view and the
text view.

We often speak as if properties have parts, yet none of the available conceptions captures this
way of speaking. None of the available conceptions of property parthood explicitly captures the
judgment that being mammalian is part of being a dog. None of the available conceptions of property
parthood were meant to capture this way of speaking either. But if properties have parts in the way we
often say they do, and if we have some grasp on what this means, we can leverage that understanding
to propose a set of axioms which implicitly define property parthood.

2.5 The Mereology

In this section, I offer a formal theory of property parthood. I’ll begin with definitions. In what
follows, lower-case variables (e.g., \( x, y, z \)) range over properties instead of objects.\(^{23}\) Furthermore,
we’ll define other mereological notions in terms of ‘proper part’, our chosen primitive.\(^{24}\) The
definitions straightforwardly carry over from CEM.

**Definition 1.** \( x \) is part of \( y \) iff \( x \) is a proper part of \( y \) or identical to \( y \).

**Definition 2.** \( x \) and \( y \) overlap iff there is some \( z \) which is part of both \( x \) and \( y \).

These definitions are standard fare in mereology. Applying them to properties is not, however, so
let’s detour through some examples. Being mammalian is part of being a cat in the same sense that
being red is part of being crimson. But if being a cat were part of being mammalian, all mammals
would be cats. Some mammals are not cats, of course, so being a cat isn’t part of being mammalian.
Thus, being mammalian and being a cat aren’t identical, by the non-identity of discernibles. Being
mammalian is part of but not identical to being a cat, so by Definition 1, being mammalian is a
proper part of being a cat.

\(^{23}\)I will also use upper-case variables occasionally to highlight the link between a predicate ‘\( F \)’ and the property of being
\( F \). I eschew upper case variables in the formal presentation to avoid confusion between the standard meaning of plural
referring expressions of the form ‘the Fs’ (the things that are \( F \)) and plural variables that range over properties, which, if I
used upper case variables, would also be of the form ‘the Fs’.

\(^{24}\)They are many equally legitimate candidate notions to choose as a primitive. For instance, we might have chosen ‘part’
as our primitive and said that \( x \) is proper part of \( y \) iff \( x \) is a part of \( y \) and not identical to \( y \). The rest would look the same.
For similar reasons, being mammalian is also a proper part of being a dog. Being a mammal is part of what it takes to be a dog and also part of what it takes to be a cat. Being mammalian, then, is part of both being a cat and being a dog, which means that being a dog and being a cat overlap.

Using these definitions, we define two other notions:

**Definition 3.** $x$ and $y$ are disjoint iff $x$ and $y$ do not overlap.

**Definition 4.** $x$ is a sum of the $y$s iff each of the $y$s is part of $x$ and any $z$ overlaps one of the $y$s iff $z$ overlaps $x$.

Definition 3 introduces the idea of disjointness between properties. Whether there are any disjoint properties is disputable, and I’ll explain this issue shortly. Definition 4 introduces the idea of a sum of properties as itself a property. Roughly, a sum includes every property which is part of what it takes to have that property. Being the actual Wayne Gretzky, for example, is a sum of properties whose parts are the properties which correspond to everything truly predicable of Gretzky. Which other properties are parts of being (actual) Wayne Gretzky will depend on whether we adopt a coarse- or fine-grained account of properties. A more coarse-grained, modal conception of parthood, which I reject, may imply that being such that $2 + 2 = 4$ is part of what it takes to be the actual Wayne Gretzky. On a more fine-grained approach, this will not follow. But adopting one or the other is compatible with this standard definition of mereological summation.

The first axiom says that proper parthood is asymmetric:

**Asymmetry.** If $x$ is a proper part of $y$, $y$ isn’t a proper part of $x$.

The motivation for Asymmetry is that nothing can subsume and outstrip something that simultaneously and in the same sense subsumes and outstrips it. Properties subsume and outstrip their proper parts, so no property is a proper part of any of its proper parts. Let’s return to the properties being mammalian and being a cat. Being mammalian is a proper part of being a cat because there’s more to being a cat than to being a mammal. But being a cat isn’t a proper part of being mammalian. If it were, we could infer from Definition 1 that being a cat is part of being mammalian. Being a cat isn’t part of being mammalian because some mammals aren’t cats. Since being a cat isn’t part of being mammalian, it isn’t a proper part of being mammalian either, by Definition 1.
The following principle follows from Definition 1 and Asymmetry:

**Uniqueness.** \( x \) and \( y \) completely overlap iff \( x \) and \( y \) are identical.

If \( x \) and \( y \) are identical, we can exchange the terms referring to them in any statement about either’s having a part *salva veritate*. So if \( x \) and \( y \) are identical, whatever properties are truly said to be parts of one are also truly said to be parts of the other. So identity makes for complete overlap. This secures the right-to-left portion of Uniqueness. Definition 1 implies that each property is part of itself. If \( x \) is part of itself, \( y \) is part of itself, and \( x \) and \( y \) completely overlap, then \( x \) is part of \( y \) and \( y \) is part of \( x \). From Definition 1 we can then infer that \( x \) and \( y \) are either proper parts of each other or identical. They’re identical, then, because Asymmetry prohibits their being proper parts of each other. This secures the left-to-right portion of Uniqueness.

The second axiom says that proper parthood is transitive:

**Transitivity.** If \( x \) is a proper part of \( y \), and \( y \) is a proper part of \( z \), then \( x \) is a proper part of \( z \).

To illustrate this axiom, consider the trio of properties, *being an animal*, *being mammalian*, and *being a dog*. *Being an animal* is a proper part of *being mammalian* because *being an animal* is part of but not identical to *being mammalian*. And *being mammalian* is a proper part of *being a dog* because *being mammalian* is part of but not identical to *being a dog*. We’ve satisfied the antecedent of Transitivity, so it follows that *being an animal* is part of *being a dog* (but not vice versa). So *being an animal* is a proper part of *being a dog*. This is intuitively correct.

Now let’s consider two principles whose axiomatic status isn’t as clear. Here is the first:

**Weak Supplementation.** If \( x \) is a proper part of \( y \), then \( y \) has another proper part disjoint from \( x \).

The principle’s name is meant to suggest that every proper part accompanies another disjoint proper part. Whether Weak Supplementation holds is not clear, a status it shares with its analogue for spatio-temporal parthood.\(^{25}\) Weak Supplementation rules out the scenario in which a property has a single proper part, as well as the scenario in which each of a property’s proper parts overlaps every

\(^{25}\)See, for example, Smith (2009).
other. On the modal analysis of parthood (being $F$ is part of being $G$ if and only if, necessarily, whatever is $G$ is $F$) Weak Supplementation fails universally. Given any two properties being $F$ and being $G$, there is a disjunctive property of being $F$ or $G$. Since, necessarily, whatever is $F$ is also $F$ or $G$, and, necessarily, whatever is $G$ is also $F$ or $G$, the modal analysis implies that being $F$ or $G$ is part of both being $F$ and being $G$. Indeed, the modal analysis implies that for any group of properties whatsoever, there is some disjunctive property that is part of them all.

The above argument against Weak Supplementation depends on the modal analysis of parthood. Since I reject the modal analysis for independent reasons (see Section 3), I have one less reason to deny Weak Supplementation. There may be other reasons to reject Weak Supplementation, but very little in the following chapters hangs on its truth. In any case, I’m happy to accept Weak Supplementation and avoid the complications which result from rejecting it while simultaneously trying to save the intuition that subtracting a property’s proper part should leave some remainder.26

The following principle is also disputable:

**General Sum Principle.** For any properties whatever, there is a sum of those properties.

Let’s say, first, that two properties preclude each other when one property is the negative correlate of the other, like being not-$F$ and being $F$, and, second, that a property is inconsistent when two of its parts preclude each other. The General Sum Principle posits inconsistent properties, which some may find worrisome. Inconsistent properties, if exemplified, would pull reality in opposite directions. Reality cannot be pulled in opposite directions, so inconsistent properties remain necessarily unexemplified. What good is a property if nothing could ever exemplify it? It would be an ontological free-loader, a property which never punches it’s ticket as a property of something.

Given a Platonistic view of abundant properties, however, the General Sum principle doesn’t make us any worse off. Pick some arbitrary number of random properties, (say), being $F$, being $G$, and being $H$. The General Sum Principle says that there’s a sum composed of them. The abundant conception says that the predicate ‘is an $F$, $G$, and $H$’ expresses a property. The sums secured by the General Sum principle are the properties licensed by the abundant conception. This is no different in the case of inconsistent predicates and the properties they express. The abundant conception licenses inconsistent properties because there are inconsistent predicates. Adopting the General

26See Varzi (2009, Sec. 3.1).
Sum Principle merely allows us to identify these inconsistent properties as sums of their parts. The abundant conception says that the predicate ‘is F and not-F’, for example, expresses a property, and the General Sum Principle says that there’s a sum composed of the properties being F and being not-F. The sums secured by the General Sum principle are the properties licensed by a Platonistic abundant conception of properties. The General Sum Principle commits the Platonist about abundant properties to no more properties than she already was.

The final two axioms govern properties of properties, or meta-properties. The property of being a dog is a property of dogs, and the property of being a desk is a property of desks. These are properties objects have. There are also properties that properties have. Whereas the property of being a dog is a property of dogs, the property of being the property of being a dog is a property’s property, a meta-property. The abundant conception licenses these meta-properties: we have the predicate ‘is the property of being a dog’ and so there’s a corresponding second-order property. These properties, too, have parts.

Here is the first axiom for meta-properties:

**Inclusivity:** If $x$ is part of $y$, then having $x$ as a part is part of the property of being $y$.

*Being mammalian* is part of being a dog. And what it is to be the property of being a dog includes its having being mammalian as a part. *Having the property of being mammalian as a part* is part of being the property of being a dog. More generally, Inclusivity says that some parts of meta-properties concern which properties are parts of their base properties. Being a base property or meta-property is a relative matter, for we can treat any meta-property as a base property. For instance, we can treat being the property of being a dog as the base property of being the property of being the property of being a dog.

The final axiom closely resembles Inclusivity:

**Exclusivity:** If $x$ isn’t part of $y$, then the property of not having $x$ as a part is itself part the property of being $y$.

Exclusivity says that if a property is not part of another property, the latter’s not having that property as a part is part of being that latter property. That is, some parts of meta-properties concern which
properties are not parts of their base properties. *Being a dog*, for example, is not part of *being human*. So the property of not having *being a dog* as a part is itself part of the property of *being human*.

With the presentation of the axiomatic system now complete, we can see when and how it differs from CEM. This property mereology resembles CEM in some important respects. The definitions of mereological notions are analogous, and the property mereology also contains an analogue of CEM’s General Sum Principle. There are also important differences. It has axioms with no analogues in CEM, namely, Inclusivity and Exclusivity. There are no analogues for these in CEM because there aren’t such things as meta-objects, objects of objects.

2.6 Conclusion

The positions I develop in the following chapters all depend in some way on the legitimacy of property parthood. Those positions require intensional entities, whether we call them “properties,” or “concepts” or something else. They also require that those intensional entities have the sort of hierarchical structure we often attribute to them in our property parthood judgments. So they require that being an animal is part of being a dog, for instance, and not the other way around. Finally, the positions require that some though probably not all of the mereological axioms hold. Some of these axioms are more central to the project than others. For instance, the main positions might very well survive a rejection of Weak Supplementation. But I doubt the same could be said of Transitivity.

My rejection of the modal analysis in Section 3 raises the question of how fine-grained the present intensional entities are. Consider, for instance, the property of being a woodchuck and the property of being a groundhog.27 In some sense, one may consistently believe that Tom is a woodchuck without believing that Tom is a groundhog. If properties serve as the semantic values of ‘is a woodchuck’ and ‘is a groundhog’ and the semantic values of predicates should explain why one may, in some sense, consistently believe that something is a woodchuck and not believe (or disbelieve) that the same thing is not a groundhog, then some may insist on a more fine-grained notion of property which distinguishes the property of being a groundhog from the property of being a woodchuck.

The view I’ve offered here does not permit us to distinguish the property of being a woodchuck from the property of being a groundhog in that way. To be a groundhog is to be a woodchuck and vice versa. The property of being a woodchuck is identical with the property of being a groundhog for they have all the same parts. These properties are identical whether we believe it or not. Yet I later offer these properties as the meanings of predicates. How could we have some understanding of the meaning of two predicates without realizing that they have the same meaning? We may consistently believe that something is a woodchuck but not a groundhog because we attach one word to an insufficient or incorrect understanding of the property of being a groundhog/woodchuck and attach another word to another possibly insufficient or mistaken understanding of the property. We do so without realizing that the property we have some insufficient or incorrect understanding of is the very same property we have some other possibly insufficient or incorrect understanding of.

The property mereology I’ve offered here attempts to capture many of our ordinary parthood judgments. Although the property mereology is intrinsically interesting for that reason, it is also theoretically useful. In the following three chapters, I use the account of property parthood here to develop novel semantic approaches to first-order logic, modal propositional logic, and quantified modal logic. I argue that in some ways these approaches work better than the standard approaches. I then use it in accounts of sets and the natural numbers. It is a virtue of the mereology that it captures a way we often talk about properties. But it is also a virtue that we can use it for so many other purposes.
CHAPTER 3

Modal Intensionalism

3.1 Introduction

Contemporary discussions of necessity and possibility revolve around the view that our world is one among many possible worlds. On the standard possible worlds approach to modality, a proposition is necessarily true when it is true in all possible worlds and possibly true when it is true in at least one. Possible worlds have been very useful, but there is no consensus about what they are or even whether they are.

Broadly speaking, two realist views of modality have been highly influential. According to David Lewis’s modal realism, possible worlds are spatio-temporally isolated island universes. Any way things could be is the way some island universe really is. Necessarily true propositions are true in every island universe, and possibly true propositions are true in at least one island universe. Although each possible world is a full-blooded universe, only one is actual. For Lewis, ‘actual’ is an indexical like ‘here’: just as here’ refers to the place of utterance and to no other places, ‘actual’ refers to the island universe of utterance and to no others.

Ersatzism foregoes Lewis’s concrete universes and associates possible worlds with abstract representations of ways the world might have been. Following Lewis, I’ll call an abstract object (or collection of abstract objects) a world-surrogate if it represents a complete way a world might have been. According to ersatzism, necessarily true propositions are those which each world-surrogate represents as being so, and possibly true propositions are those which at least one world-surrogate represents as being so. World-surrogates are variously identified with consistent sets of propositions,

---

1 Lewis (1986b).
2 Lewis (1986b, Cf. 137).
exemplifiable properties, and obtainable states of affairs.\(^3\) Since these entities exist in the actual world, ersatzers need not endorse Lewis’s indexical theory of actuality.

Whatever possible worlds are, we often do not appeal to them when we explain why things must be a certain way. Instead of saying that any mammal must be an animal because, in every possible world, every mammal is also an animal, we might say that any mammal must be an animal because being an animal is part of being mammalian. We seem to think that nothing could exemplify the whole property without exemplifying its parts, and we associate the parts with preconditions for exemplifying the whole.\(^4\) If we take this way of speaking seriously, an entirely new semantics for modal logic results. Instead of treating necessary truths as truths in every possible world, we may treat them as the preconditions for the existence of a world in general. In this paper, I develop this alternative approach to modality for modal propositional logic. The approach is not Kripkean in any way: neither possible worlds nor accessibility relations nor formally similar stand-ins of either appear in the semantic machinery.

The machinery revolves around two properties. First, there’s the property of being our world. It’s part of being our world—call it \(alpha\)—that Caesar crosses the Rubicon. I take this “part”-talk literally and use it to treat a proposition’s truth-value: a proposition \(\phi\) is true just in case its corresponding propositional property \(\text{being such that } \phi\) is part of \(\text{being alpha}\).\(^5\) Secondly, we can abstract away from the property of being our world to the property of being a world in general. When there is a precondition for being a world, its being the case is part of \(\text{being a world}\). When its being the case is part of \(\text{being a world}\), it’s necessary. Once more, I take this “part”-talk literally: a proposition \(\phi\) is necessarily true just in case \(\text{being such that } \phi\) is part of \(\text{being a world}\).

My approach is committed to an ontology of properties and to the intelligibility of a parthood relation among those properties. I discuss these commitments in Chapter 1. After I explicate the approach to modality in Section 3, I show in Section 4 that principles inspired by that property

---

3See, for example, Adams (1974), Stalnaker (1976), and Plantinga (1974).

4Leibniz presents a similar view of the part-whole structure of concepts (See C 53, 55 in Couturat (1903) and its translation in Parkinson (1966) at P 20, 22.) Leibniz associates a concept’s parts with its “requisites,” and claims that the concept of being metal is part of the concept of being gold because all the requisites of being metal are among the requisites of being gold.

5I will ignore the standard markers for use and mention when the meaning is clear.
mereology validate axioms of the normal modal propositional calculi. In Section 6, I argue that, compared to possible worlds semantics, the meanings in the new semantics operate on a more basic level of modal reality, both metaphysically and epistemologically. In the next two sections, I highlight a feature embedded in the formalism of possible worlds semantics and show how embedding another feature yields the non-Kripkean alternative.

3.2 Inclusion

My approach and the possible worlds approach treat modal space differently. To bring out the difference, consider the sentence ‘Fred is tall’. Standard treatments say that this sentence is true just in case Fred is in the extension of ‘tall’, the class or set of things which are tall. They treat the truth-value of a subject-predicate sentence as if it depends on whether the subject’s referent is a member of the predicate’s extension. Membership in a set or class is an extensional inclusion relation. So call the approach to logic which treats a sentence’s truth-value as depending on extensional inclusion relations logical extensionalism.

The traditional alternative to logical extensionalism appeals to a relation between intensional entities. Intensional entities are associated with extensions in the actual world, but different intensional entities may have the very same extension. The property of being a cordate and the property of being a renate are two examples of an intensional entity. Although the class of cordates is identical to the class of renates (because they have the same members), the property of being a cordate differs from the property of being a renate.

Terms like ‘Fred’ or ‘H₂O’ may have two kinds of associated intensional entities. Usually, fine-grained intensions are identified with concepts. The fine-grained intension of ‘water’ is considered distinct from the fine-grained intension of ‘H₂O’ because we can imagine watery stuff that’s not composed of H₂O. Usually, coarse-grained intensions are identified with properties. The coarse-grained intensions of ‘water’ and ‘H₂O’ are considered identical because the properties being water and being H₂O are identical.

---

6 I borrow the terms “extensional inclusion” and “intensional inclusion” from Swoyer (1995).
A few stipulations will simplify the presentation of the alternative to logical extensionalism. First, I will frame the ensuing discussion of intensional entities in terms of “properties” but I will leave open how fine- or coarse-grained they are, at least insofar as my approach allows. Second, I endorse an abundant conception of properties, which I explain in Chapter 1. Finally, I assume that predicates express properties. So I will say that the predicate ‘is tall’, for example, expresses the property of being tall.

The alternative to logical extensionalism does not treat a sentence’s truth-value as depending on extensional inclusion relations but instead on whether the subject’s intension, a property, includes the property that the predicate expresses. I’ll call the inclusion relation between properties intensional inclusion. On this approach, the sentence ‘Fred is tall’ is true just in case the property of being Fred includes the property of being tall. For the obvious reason, I’ll call the approach to logic which treats a sentence’s truth-value as depending on intensional inclusion relations logical intensionalism.

The crucial difference between logical extensionalism and logical intensionalism is the direction of their respective inclusion relations. Whereas the set of tall things includes Fred on the extensionalist approach, being Fred includes being tall on the intensional approach. Leibniz knew of the inverse relationship between the approaches and famously favored an intensional approach.⁸

I’ve defined intensional approaches in terms of intensional inclusion. But it is often more natural to speak of a property’s parts rather than the properties a property includes. Since property inclusion is the converse of property parthood (being F is part of being G just in case being G includes being F), nothing is lost if we construe intensional approaches in terms of property parthood. For most of what follows, I will opt for property parthood, which I discuss in Chapter 1. In the next section, I show how a new kind of semantics results when we replace the extensional inclusion relation embedded within possible worlds semantics with an intensional inclusion relation.

### 3.3 Frames

A semantics for a modal logic may accomplish one or two overarching tasks. Minimally, a semantics for modal logic uses a model structure, a set-theoretical construction, to determine the truth conditions for formulas, including formulas with modal operators like □ and ◊. On the basis of those truth

---
⁸See C 53/P 20.
conditions, this machinery determines which formulas are valid in which systems. But a semantics may do this and leave the modal operators uninterpreted, disconnected from any modal notions. Alvin Plantinga (1974, 127) aptly calls such a semantics pure. Kripke semantics is pure and doesn’t have any explicit connection to the modal notions of necessity and possibility.\(^9\) A model structure in Kripke semantics consists of an ordered triple \(\langle G, K, R \rangle\), where \(G\) is a member of \(K\) and \(R\) is a binary relation defined over the members of \(K\). I’ll assume some familiarity with how these function in Kripke semantics.

One may also intend an interpretation of the model structure to model some portion of modal discourse. Possible worlds semantics co-opts the formalism of Kripke semantics and explicitly associates the members of \(K\) with possible worlds (or world-surrogates), \(G\) with the actual world, and \(R\) with an accessibility relation defined over those possible worlds. One may use possible worlds semantics to model metaphysical, deontic, epistemic or other kinds of modality with different readings of the modal operators. For example, the box may read “it is obligatory that” in deontic logic or “it is known that” in epistemic logic. The possible worlds approach to metaphysical modality says that \(\Box \phi\) (read “necessarily, \(\phi\)”) is true at a world just in case \(\phi\) is true in every possible world accessible from it and that \(\Diamond \phi\) (read “possibly, \(\phi\)”) is true at a world just in case \(\phi\) is true in some accessible possible world.

My approach and the possible worlds approach to metaphysical modality differ in their intended model structures—what I’ll henceforth call frames.\(^{10}\) The frames differ in whether the embedded relation between actuality and worldhood is an intensional or extensional inclusion relation. The possible worlds approach to metaphysical modality has a set of possible worlds, \(K\), one of which is actual—\(G\). This is an extensional inclusion relation. Call the possible worlds approach to metaphysical modality where the actual world (or its surrogate) is a member of a certain class of possible worlds (or their surrogates) modal extensionalism.

The alternative inverts the relationship between actuality and worldhood with an intensional inclusion relation. Call the world in which we live alpha. Modal intensionalism is the semantic approach to metaphysical modality according to which being a world is part of being alpha. An

---


\(^{10}\) A “frame” in possible worlds semantics is usually an ordered pair of a set of worlds and an accessibility relation without any world pegged as actual.
intensionalist frame consists of an ordered triple, \( (A, W, P) \), where \( A \) is the property of being our world \textit{alpha}, \( W \) is the property \textit{being a world}, and \( P \) is a parthood relation defined over properties. The diagram below depicts the two types of frames:

![Diagram of modal intensionalism and modal extensionalism frames]

The modal extensionalist’s frame is the modal intensionalist’s turned “inside out,” and this difference unsurprisingly yields different treatments of the modal notions. Whereas modal extensionalism says that a proposition \( \phi \) is necessarily true when it is true in every possible world (or world-surrogate), modal intensionalism says that a proposition \( \phi \) is necessarily true when the property of being such that \( \phi \) is part of \textit{being a world}.

In what follows, I develop modal intensionalism for modal propositional logic. Since, among the standard modal systems, many believe the S5 system captures metaphysical modality best, I will interpret the box as “it is metaphysically necessary that” and provide an applied semantics for the S5 system of modal propositional logic.

### 3.4 Truth, Necessity, and Possibility

An intensionalist approach to any logic requires that intensional inclusion relations among properties determine the assignment of truth-values. But an intensionalist approach to the modal propositional calculi cannot appeal to properties associated with subjects or predicates because there are no subjects and predicates in the propositional calculus. So intensional inclusion relations among properties may determine the assignment of truth-values to propositions only if the properties correspond to the unaanalyzed propositions themselves.
There’s a function from propositions to the relevant properties. Start with the usual stock of proposition letters \( p_0, p_1, p_2, \ldots \), which represent propositions not further analyzed. Modal propositional logic contains non-atomic propositions, too, and those non-atomic propositions require connectives. So we’ll throw in the connectives written \( \neg, \land, \lor, \text{ and } \supset \), which represent negation, conjunction, disjunction, and the material conditional, respectively. Now let ‘[...]’ be a 1-1 function from propositions to their corresponding propositional properties. So for every proposition \( \phi \), there is a property \([\phi]\), or being such that \( \phi \). Given this function, \([\phi]\) is identical to \([\psi]\) just in case \( \phi \) is identical to \( \psi \).\(^{11}\)

There are at least two ways in which propositional properties differ from non-propositional properties. When something exemplifies a non-propositional property, the non-propositional property is part of being that thing. For example, when Fred is tall, being tall is part of being Fred. But at least in simpler cases, when something exemplifies a propositional property, it is part of being that thing that some property is part of another. When alpha exemplifies the propositional property of being such that Fred is tall, for instance, it is part of being alpha that being tall is part of being Fred.

Secondly, propositional properties have a special connection to actuality and truth because of their additional structure. On the intensionalist approach, a proposition \( \phi \) is true just in case its corresponding property being such that \( \phi \) is part of being alpha.\(^{12}\) So on the intensionalist approach, there is an equivalence between truth and actuality.\(^{13}\) A propositional property’s being part of being alpha and its corresponding proposition’s being true are one and the same. Where \( A \) is the property being alpha and ‘\(<\)’ reads “is part of”:

\[
(A) \quad \phi \text{ is true } =_{df} [\phi] < A.
\]

Here’s how this treatment “looks”:

Since the S5 system contains classical logic, and since I aim to offer an applied semantics for S5, I will assume the world behaves classically in ways that secure the theorems of classical logic. So I assume that \( A \), the property of being alpha, is complete: for every pair of propositions \( \phi \) and \( \neg \phi \).

\(^{11}\) Zalta (1993, 279, n. 7)

\(^{12}\) Most think that truth is “extensional” in the sense that when modal operators are not an issue, one may substitute any true proposition for any other \( \text{salva veritate} \). This is compatible with the present intensionalist treatment of truth.

\(^{13}\) This is the intensionalist side of the “true-story” theory of actuality in Adams (1974, 226).
at least one of either \([\phi]\) or \([\neg\phi]\) is part of \(A\). I also assume that \(A\) is consistent: for every pair of propositions \(\phi\) and \(\neg\phi\), at most one of either \([\phi]\) or \([\neg\phi]\) is part of \(A\). Therefore, if \([\phi]\) is a part of \(A\), \([\neg\phi]\) isn’t, and vice versa. \(A\)’s completeness secures the law of excluded middle and guarantees that there are no truth-value gaps. Its consistency secures the law of non-contradiction and guarantees that there are no true contradictions. One might reject \(A\)’s completeness or consistency because of vagueness or liar paradoxes, but I ignore these complications because the first test for an alternative approach to modal logic should be whether and how it handles the classical systems.

A number of auxiliary assumptions seem to follow from our intuitive understanding of ‘...and...’ statements, ‘...or...’ statements, ‘if... then...’ statements and the like. For example:

(i) if \([\phi \supset \psi]\) and \([\phi]\) are parts of \(A\), so is \([\psi]\),

(ii) \([\phi]\) is part of \(A\) iff \([\neg\neg\phi]\) is,

(iii) \([\phi \land \psi]\) is part of \(A\) iff \([\phi]\) and \([\psi]\) are, and

(iv) \([\phi \lor \psi]\) is part of \(A\) iff either \([\phi]\) or \([\psi]\) is.

These secure theorems and validate various inferences of propositional logic. For example, if \([\phi \land \psi]\) is part of \(A\), then the conjunction \(\phi \land \psi\) is true, by (A). But \([\phi \land \psi]\) is part of \(A\) if and only if both \([\phi]\) and \([\psi]\) are, by (iii). So if \([\phi \land \psi]\) is part of \(A\), then so are \([\phi]\) and \([\psi]\). If \([\phi]\) and \([\psi]\) are both parts of \(A\), then both \(\phi\) and \(\psi\) are true, by (A). Hence, (iii) validates the inferences from a conjunction to each conjunct.

Now we proceed from truth to necessary truth. Modal extensionalism treats necessary truth as truth in every possible world:
But modal intensionalism says that $\phi$ is necessarily true if and only if its corresponding propositional property *being such that* $\phi$ is part of *being a world*. More formally, where $W$ is the property *being a world*:

$$(N) \quad \Box \phi \text{ is true} \iff [\phi] < W.$$ 

In the S5 system, we may depict (N) as follows:

Like the propositional parts of $A$, I assume (reasonably, I think) that the propositional parts of $W$ behave in ways that secure our intuitive understanding of ‘...and...’ statements, ‘...or...’ statements, ‘if... then...’ statements and so on. This suggests a number of principles that govern how propositional properties operate in $W$, including:

(v) if $[\phi \supset \psi]$ and $[\phi]$ are parts of $W$, so is $[\psi]$,

(vi) $[\phi]$ is part of $W$ iff $[\neg \neg \phi]$ is,
(vii) \([\phi \land \psi]\) is part of \(W\) iff \([\phi]\) and \([\psi]\) are, and

(viii) if \([\phi]\) is part of \(W\), so is \([\phi \lor \psi]\).\(^{14}\)

Joined with (N), these principles validate a number of intuitive modal inferences. Consider (vii), for instance. It can be easily shown that (vii) and (N) validate the inferences from \(\Box(\phi \land \psi)\) to both \(\Box \phi\) and \(\Box \psi\) and from both of these back again to \(\Box(\phi \land \psi)\).

These sorts of principles, along with (N), imply that the theorems of propositional logic are necessarily true. For example, given (vii), if \([\phi \land \psi]\) is part of \(W\), \([\phi]\) is, too. Then, based on our intuitive understanding of ‘if... then...', \(((\phi \land \psi) \supset \phi)\) is also part of \(W\). We then infer from (N) that \((\phi \land \psi) \supset \phi\) is necessarily true. The property \(((\phi \land \psi) \supset \phi)\) corresponds to a theorem of propositional logic. Intuitively, all other theorems of propositional logic have corresponding propositional properties that are also parts of \(W\), so they are also necessarily true. These considerations justify an important inference rule in the weakest normal modal system K (a system that has all the theorems of propositional logic as theorems):

*\textbf{Necessitation Rule.} If \(\phi\) is a theorem of K, then so is \(\Box \phi\).*

Next we define possibility in terms of necessity. \(\neg \phi\) is necessary just in case \(\phi\) is impossible. So \(\phi\) is not impossible—i.e., possible—just in case \(\neg \phi\) is not necessary. On the intensionalist approach, then, a proposition \(\phi\) is possible when \(\neg \neg \phi\) is part of \(W\):

\[\text{(P) } \Diamond \phi \text{ is true } =_{df} \neg \neg \phi \notin W\]

Hence, what is possibly true corresponds to what \(W\)'s parts do not preclude. Traditionally understood, the necessity and possibility operators are interdefinable: \(\Box \phi\) is equivalent to \(\neg \Diamond \neg \phi\) (and \(\neg \Box \neg \phi\) is equivalent to \(\Diamond \phi\)). Modal intensionalism justifies this equivalence. First, suppose that \(\neg \Diamond \neg \phi\) is true. By (P), \(\Diamond \neg \phi\) is true when \(\neg \phi\) is not part of \(W\). So \(\neg \Diamond \neg \phi\) is true when \(\neg \phi\) is part of \(W\). Given that \(\neg \phi\) is part of \(W\), \([\phi]\) is part of \(W\), too, by principle (vi). As a result, \(\Box \phi\) is true, by (N). Therefore, if \(\neg \phi\) is true, then so is \(\Box \phi\). Now suppose that \(\Box \phi\) is true. By (N), \([\phi]\) is part of \(W\). Again, by principle (vi), since \([\phi]\) is part of \(W\), \(\neg \neg \phi\) is also part of \(W\). And, as before, \(\neg \Diamond \neg \phi\) is true when \(\neg \neg \phi\) is part of \(W\). Therefore, if \(\Box \phi\) is true, \(\neg \Diamond \neg \phi\) is true, which completes the proof.

\(^{14}\) An analogue of (iv), the stronger biconditional principal that \([\phi \lor \psi]\) is part of \(W\) iff either \([\phi]\) or \([\psi]\) is, presumably doesn’t hold for \(W\) because \(W\) isn’t complete like \(A\) is.
3.5 Pure Semantics

If modal intensionalism’s formalism has anything like the expressive power of possible worlds semantics, different restrictions on its formal frames should validate different modal formulas. In this section, I ask the reader to leave \( \mathcal{A} \) and \( \mathcal{W} \) uninterpreted and construe \( \mathcal{P} \) as an attenuated notion of parthood. With this attenuated notion, we may say that a “property” is “part” of another without assuming that the relation obeys the axioms or theorems of the property mereology from Chapter 1. I will show how formal restrictions on \( \mathcal{P} \) validate the characteristic axiom schemata of some normal modal systems. For each axiom schema, we place a restriction on \( \mathcal{P} \) and abstract from the other potential restrictions. This is an exercise in pure semantics. In the next section, I’ll interpret \( \mathcal{P} \) as the relation of property parthood in an applied semantics for metaphysical modality.

Principle (v) validates the (K) axiom:\(^{15}\)

\[ \mathcal{K} \quad \Box(\phi \supset \psi) \supset (\Box\phi \supset \Box\psi) \]

For \( \Box(\phi \supset \psi) \) is true just in case the property \([\phi \supset \psi]\) is part of \( \mathcal{W} \). Now suppose that \([\phi]\) is part of \( \mathcal{W} \) (i.e., that \( \Box\phi \) is true). Then \([\psi]\) is also part of \( \mathcal{W} \), by (v). As a result, \( \Box\psi \) is true, by (N). So (K) is valid when \( \mathcal{P} \) obeys (v).

A property is consistent only if at most one of either \([\phi]\) or \([\neg\phi]\) is part of it. \( \mathcal{W} \)’s consistency secures the following:

\[ \mathcal{D} \quad \Box\phi \supset \Diamond\phi \]

Given \( \mathcal{W} \)’s consistency, if \([\phi]\) is part of \( \mathcal{W} \) (i.e., if \( \Box\phi \) is true, by (N)), then \([\neg\phi]\) is not part of \( \mathcal{W} \) (i.e., \( \Diamond\phi \) is true, by (P)). So (D) is valid on frames in which \( \mathcal{W} \) is consistent.

I’ll say that \( \mathcal{P} \) is connected when \([\phi]\) is part of \( \mathcal{A} \) if \([\phi]\) is part of \( \mathcal{W} \). Connectedness is associated with the following:

\[ \mathcal{T} \quad \Box\phi \supset \phi \]

\(^{15}\)Strictly speaking, (K) is an axiom schema whose instances are axioms in the K-system. I’ll often talk as if various axiom schemata are axioms, as if a schema is valid when all of its instances are, and as if a schema is true when its instances are all true.
Suppose that $[\phi]$ is part of $\mathcal{W}$. Given $\mathcal{P}$’s connectedness, $[\phi]$ is then part of $\mathcal{A}$. Hence, if $[\phi]$ is part of $\mathcal{W}$ and $\Box \phi$ is true (by (N)), then $[\phi]$ is part of $\mathcal{A}$ and $\phi$ is true (by (A)). So (T) is valid on connected frames.

The next three restrictions are new, and I will explain them further in the next section. The first of these requires a definition: a property is *ininclusive* just in case whenever it has a property as a part, the property of being such that it has that property as a part is itself a part of that property. For example, if $\mathcal{W}$ is ininclusive and $[\phi]$ is part of it, then being such that $[\phi]$ is part of $\mathcal{W}$ is itself part of $\mathcal{W}$. If $\mathcal{W}$ is ininclusive, the following axiom schema is valid:

\[(S4) \Box \phi \supset \Box \Box \phi.\]

Suppose that $\mathcal{W}$ is ininclusive in $\mathcal{P}$. If $[\phi]$ is part of $\mathcal{W}$, so is the property of being such that $[\phi]$ is part of $\mathcal{W}$, written $[[\phi] \upharpoonright \mathcal{W}]$. According to (N), if $[\phi]$ is part of $\mathcal{W}$, $\Box \phi$ is true. So if $[[\phi] \upharpoonright \mathcal{W}]$ is itself part of $\mathcal{W}$, then, in a sense, $\phi$’s being necessary is itself part of $\mathcal{W}$. (N) kicks in once more and we conclude that $\Box \Box \phi$ is true. (S4) is valid on ininclusive frames (i.e., frames where $\mathcal{W}$ is ininclusive).

Let’s call the following condition on $\mathcal{P}$ *fortification*: if $[\phi]$ is part of $\mathcal{A}$ then $[-\phi]$’s not being part of $\mathcal{W}$ is itself part of $\mathcal{W}$. When $\mathcal{P}$ is fortified, the following schema is valid:

\[(B) \phi \supset \Box \Diamond \phi\]

Suppose $[\phi]$ is part of $\mathcal{A}$. By (A), $\phi$ is true. Then, by fortification, $[-\phi]$’s not being part of $\mathcal{W}$ is part of $\mathcal{W}$. By (P), $\phi$’s being possible is part of $\mathcal{W}$. And then by (N), $\Box \Diamond \phi$ is true. Hence, (B) is valid on fortified frames.

Finally, call a property *inexclusive* when it meets the following condition: if a property is not part of it, then being such that the property is not part of it is itself part of it. If $\mathcal{W}$ is inexclusive, then if $[\phi]$ is not part of $\mathcal{W}$, the property of being such that $[\phi]$ is not part of $\mathcal{W}$ is itself part of $\mathcal{W}$. The axiom schema below is valid when $\mathcal{W}$ is inexclusive in $\mathcal{P}$:

\[(S5) \Diamond \phi \supset \Box \Diamond \phi\]

Let’s suppose that $\mathcal{P}$ is inexclusive (i.e., that $\mathcal{W}$ is inexclusive in $\mathcal{P}$). If $[-\phi]$ is not part of $\mathcal{W}$, then $\Diamond \phi$ is true, by (P). This is what we find in (S5)’s antecedent. Given $\mathcal{W}$’s in exclusivity, if $[-\phi]$ is
not part of $W$, then $[\neg \phi \not\in W]$ (read “being such that $\neg \phi$ is not part of $W$”) is part of $W$. And if $[\neg \phi \not\in W]$ is part of $W$, we can infer that it is part of $W$ that $\phi$ is possible, by (P). Then we infer, by (N), that $\Box \Diamond \phi$ is true. This is what we find in (S5)’s consequent. Therefore, (S5) is valid on inexclusive frames.

3.6 Applied Semantics

In the previous section, I showed how (v) validates (K), $W$’s consistency validates (D), and the connectedness, ininclusivity, fortification, and inexclusivity conditions validate the (T), (S4), (B), and (S5) axioms, respectively. When we interpret $A$ as the property of being alpha, $W$ as the property of being a world, and $P$ as the parthood relation from Chapter 1, there is good though not conclusive reason to think that S5 correctly models metaphysical modality. For as long as we assume that the world behaves classically (as I outlined in section 3), there is reason to believe that $P$ obeys all these restrictions. Below, I explain these restrictions in the context of an applied semantics for metaphysical modality.

(K) and (D) are true if Principle (v) and $W$’s consistency hold. I’ve simply assumed that (v) is true and that $W$ is consistent, and I suppose someone could reject one or the other. But these are reasonable assumptions, nonetheless. Principle (v) seems to follow from our intuitive understanding of the material conditional. For if $[\phi \supset \psi]$ and $[\phi]$ are parts of $W$, surely $[\psi]$ is. Also, our world exemplifies the property of being a world, and I cannot fathom how an exemplified property could have inconsistent propositional properties as parts.

(T) is true if $P$ is connected, i.e., if the propositional properties which are part of $W$ are also part of $A$. We have reason to think that this condition holds. First, being a world in general ($W$) is part of being this world ($A$). Second, the transitivity of property parthood is a theorem of the property mereology. Therefore, if $[\phi]$ is part of $W$, and $W$ is part of $A$, then $[\phi]$ is part of $A$. So on the intended interpretation, we now have good reason to accept (K), (D), and (T).

Let me mention a potentially curious result. As part of $A$, $W$ inherits $A$’s consistency: if there are no properties in $A$ that preclude one another, then a fortiori there are no properties in $W$ that preclude one another. Given (A), $A$’s consistency guarantees that there are no true contradictions.
Given (N), \( W \)'s consistency guarantees that there are no true contradictions necessarily. In a way, whatever secures the truth of the law of non-contradiction also seems to secure its necessary truth.

I introduced the next three restrictions in the previous section. I will explain the first of these restrictions by describing a special feature of \( A \). Take some propositional property \([p_1]\) which is part of \( A \). Since \([p_1]\) is part of \( A \), presumably some true proposition \( p_2 \) says that \([p_1]\) is part of \( A \). So from (A), we infer that \( p_2 \)'s corresponding propositional property \([p_2]\) is also part of \( A \). Therefore, if \([p_1]\) is part of \( A \), so is the property of being such that \([p_1]\) is part of \( A \). The privilege of having properties concerning which properties are parts of a property is generally reserved for meta-properties via Inclusivity. But in virtue of being a property which accounts for all truths, even truths about itself, \( A \) is ininclusive and has properties which concern which properties are parts of itself.

The propositional parts of \( W \) tell us “what it takes” for any world to exist at all. These parts of \( W \) concern the preconditions for the existence of a world in general, and I doubt whether the preconditions for worldhood could have been different. What could possibly alter the preconditions for the existence of a world in general? If the answer is “nothing,” the preconditions for anything’s being a world at all include that \( W \) has the very parts it has. If \([\phi]\) is part of \( W \), \([\phi]\) is part of \( W \) necessarily. That is, if \([\phi]\) is part of \( W \) then \([\phi]\)'s being part of \( W \) is itself part of \( W \). Like \( A \), \( W \) is ininclusive. (S4) is true if \( W \) is ininclusive.

On the current picture, \( W \) is part of \( A \) and both are consistent. Now suppose \( \phi \) is true, that \([\phi]\) is part of \( A \). Then it seems reasonable to conclude that the preconditions for worldhood could not have included that \( \phi \) is false. That is, it is a precondition that there is no such precondition that \( \phi \) be false. So not only is \([\neg \phi]\) not part of \( W \) (due to \( W \)'s being part of \( A \) and \( A \)'s consistency). Also, the property of being such that \([\neg \phi]\) is not part of \( W \) is itself part of \( W \). So if \([\phi]\) is part of \( A \), it is part of \( W \) that \([\neg \phi]\) is not part of \( W \). This restriction fortifies the truth as necessarily possible and so secures the truth of (B).

I will introduce the final restriction by describing another special feature of \( A \). There are truths about which properties are not parts of \( A \). Given (A), the propositional properties corresponding to these truths are also parts of \( A \). Suppose that \([p_1]\) is not part of \( A \). Then, it is true that \([p_1]\) is not part of \( A \). By (A), then, being such that \([p_1]\) is not part of \( A \) is itself part of being \( A \). Some of \( A \)'s parts concern which properties are not its parts. So \( A \) is not only ininclusive but also inexclusive. \( A \) itself
is a denizen of alpha, so there are truths about which properties are parts of \( \mathcal{A} \) and which are not. \( \mathcal{A} \)'s having or not having a part is part of how our world, alpha, is. In virtue of being the property of being alpha, \( \mathcal{A} \) is self-referential as both an ininclusive and inexclusive property.

\( \mathcal{W} \) is also inexclusive, though for different reasons. Earlier, I gave a reason to think that the preconditions for the existence of a world in general are preconditions necessarily. Perhaps there is also reason to think that non-preconditions for the existence of a world are non-preconditions necessarily, i.e., that \([\phi]\) must not be part of \( \mathcal{W} \) if it is not part of \( \mathcal{W} \). If \( \mathcal{W} \) is the kind of abstract object that could not have been different, then if \([\neg\phi]\) is not part of \( \mathcal{W} \), then \( \mathcal{W} \)'s not having \([\neg\phi]\) as a part is itself part of \( \mathcal{W} \). If there are necessary truths about which properties \( \mathcal{W} \) does not have as parts, then \( \mathcal{W} \) has parts that concern which properties it does not have as parts. So, like \( \mathcal{A} \), \( \mathcal{W} \) is inexclusive. (S5) is true if \( \mathcal{W} \) is inexclusive.

On the modal intensionalist’s interpretation of the formalism, one can make a decent case that the S5 system is appropriate for modeling metaphysical modality. On that interpretation, there is good reason to believe that \( \mathcal{P} \) obeys (v), the consistency of \( \mathcal{W} \), connectedness, fortification, and the ininclusivity and inexclusivity of \( \mathcal{W} \). If \( \mathcal{P} \) is restricted in all these ways, (K), (T), (D), (S4), (B), and (S5) are all true.

### 3.7 Explanation and Finite Cognition

Why should we bother with another semantics for modal logic? First, modal intensionalism can explain why the truths which happen to be true in every possible world are true in every possible world. There is reason to think, then, that modal intensionalism operates on a more fundamental layer of modal reality than modal extensionalism. Second, there are compelling reasons to think that modal intensionalism operates on a layer of modal reality at which we make many of our modal judgments, not only judgments about what is necessary or possible, but even judgments about possible worlds.

Some propositions are true in every possible world. Why are they so special? Why are they true in every possible world?\(^{16}\) The view of properties I’m exploring says that nothing could exemplify a

\(^{16}\)Michael Jubien nicely expresses the feeling that there should be some deeper explanation for why some propositions are true in every possible world. Jubien (2009, 75) writes: “the fundamental problem is that in world theory, what passes for necessity is in effect just a bunch of parallel ‘contingencies’. The theory provides no basis for understanding why these contingencies repeat unremittingly across the board (while others do not).”
property without exemplifying that property’s parts. The property of being a world is no exception. So when a proposition’s being the case is part of being a world, nothing could exemplify the property of being a world without that proposition’s being the case. A given proposition is true in each possible world because each possible world, if actual, would have exemplified the corresponding propositional property in virtue of exemplifying the property of being a world. The necessary truths are true in all possible worlds because the necessary truths correspond to the preconditions for being a world that any possible world would satisfy if it were actual. If being such that $2 + 2 = 4$ is part of being a world, for instance, then whichever possible world had been actual would have been a world (and thereby) such that $2 + 2 = 4$.

Modal extensionalism defines a proposition’s necessity as its truth in all possible worlds. Presumably, there are infinitely many possible worlds. But unless we have some heuristic, how could we know what is true in so many possible worlds? We have neither the time nor cognitive horsepower to survey so many worlds. Modal extensionalism separates the layer of modal reality which provides truth conditions for necessarily true propositions from the ways we make judgments about whether propositions are necessarily true. It separates meaning from judgment. Modal intensionalism helps reunite meaning to judgment and even supplies the heuristics to help us make judgments about infinitely many possible worlds.

Modal intensionalism provides at least two kinds of ways for us to know that something is necessarily or essentially true. First, according to modal intensionalism, we can tell whether a proposition is necessary if we can tell whether the proposition’s being the case is part of being a world. Even though being a world is presumably infinitely complex, we do not need to know about all of its parts in order to know that some property is one of its parts, especially if we grant the respectable conception of worldhood according to which a world is the totality of everything there is. On the basis of knowing that consistency is part of what it takes for there to be a totality of

---

17 Modal intensionalism provides the kind of heuristic which other proposed heuristics exploit in some way. For example, we know that a contradiction is not true in any possible world. So we may infer that the negation of the contradiction is true in all possible worlds. But how do we know that a contradiction is not true in any possible world? We seem to have some idea of what it takes to be a world. We know that consistency is part of what it takes for there to be a world.

18 This concern differs from Benacerraf-style concerns about our knowledge of abstract objects. The present concern is about the form a loosely interpreted semantics takes, not our epistemic connection to abstract objects. Granted, modal extensionalism was never meant to unite meaning and judgment. But that is all the more reason to want a modal semantics which does, not necessarily to replace modal extensionalism, but to have it at our disposal.

19 See, for example, Stalnaker (1976, 69–70).
everything there is, for example, we may know that propositional properties corresponding to logical truths are parts of being a world. Hence, we can know that many logical truths are necessary.

And if we adopt the totality conception of worldhood, we may know that propositional properties corresponding to numerical truths are parts of being a world. Being a totality of things involves facts about which things are and are not identical with one another. Facts about identity and diversity bring along with them the number that numbers the distinct things and the numbers which do not. If there is nothing but an x and a y and x is not y, there are exactly two things: numbers less than or greater than 2 do not number the things. A number numbers and the rest fail to number only if they all exist. The existence of numbers is part of what it takes for anything to exist at all. Hence, the numbers exist necessarily. From this, we may derive the necessity of arithmetical truths.

Second, the metaphysical backdrop to modal intensionalism provides another route to modal truths. Many of our modal judgments about natural kinds seem to follow this pattern. We identify some kind K, then partly on the basis of observation we build some feature F into our conception of K. We then judge that being F is part of being K to express that Ks are essentially F. We judge that being an animal is part of being mammalian, for example, partly on the basis of defeasible evidence from scientific observation. If being an animal is part of being mammalian, mammals are animals essentially. So if we have good reason to judge that being an animal is part of being mammalian, which we do, we have good reason to judge that mammals are animals essentially. We do not seem to survey infinitely many possible mammals when we make such a judgment. But the metaphysical backdrop to modal intensionalism offers a plausible, psychologically realistic story about these judgments and thus helps reunite meaning to judgment.

Modal intensionalism also provides an explanation for how we could know what is true in infinitely many possible worlds without surveying them all individually. When we make judgments about what is true in infinitely many possible worlds, we often do it in the following way. Suppose that for some proposition p, being such that p is part of being a world. Given that nothing could exemplify the property of being a world without exemplifying its parts, we may infer from the fact about parthood that p would have been true no matter which possible world had been actual just by virtue of the fact that any one of them would have been a world. Whichever one had been a

---

20This argument involves some controversial steps. I defend the connection between totality and number in Chapter 6.
world would have exemplified the property of being a world and *a fortiori* each property which is part of *being a world*, including the property of being such that *p*. If any possible world would have exemplified the property of being such that *p* if that world had been actual, then, as we say, 

‘*p* is true in all possible worlds.” Modal intensionalism supplies the heuristic to help us infer in one fell swoop what is true in all the infinitely many possible worlds. This is one way in which modal intensionalism explains the plausibility and success of modal extensionalism. This kind of explanatory support is exactly what we should expect from a semantics which operates on a metaphysically and epistemologically more basic layer of modal reality.

### 3.8 Conclusion

Modal intensionalism is an alternative approach to metaphysical modality. Like modal extensionalism, one may fill in the metaphysical details a number of ways. One might identify abundant properties with platonic universals, concepts, states of affairs, or whatever. One may even try to provide a fictionalist account of properties to cash in on the explanatory resources of modal intensionalism without committing to an ontology of properties. The plausibility of any particular intensional approach depends on how one fills in these details. There are no soundness or completeness proofs here, and modal intensionalism is decades behind modal extensionalism’s formal and philosophical successes. But I am optimistic about modal intensionalism’s prospects and its potential applications to areas of logic and philosophy which have thus far relied heavily on possible worlds.
4.1 Introduction

The two traditional semantic approaches to logic differ in their treatments of singular propositions. On the now standard approach, the singular proposition that Fred is blue is true just in case Fred is in the extension of ‘is blue’, which I’ll simply construe as the set or class of blue things.¹ This standard approach involves two commitments:

*Predicate Containment.* The semantic value of the predicate contains the semantic value of the subject.

and

*Predicate Extensionality.* The semantic value of a predicate is its extension.

I’ll call this standard approach *logical extensionalism* since it treats the truth (or falsity) of a singular proposition as if it depends on whether the predicate’s extension contains the subject’s semantic value. For convenience, I’ll often call an approach which meets these criteria an *extensional approach.*

A semantics for some discourse assigns semantic values or meanings to expressions in that discourse. To say that the extension of a predicate is its semantic value or meaning might mean nothing more than that the interpretation function assigns to each predicate its extension. However, there is a sense of ‘meaning’, along the lines of Fregean sense or Husserlian noemata, according to which the meaning of an expression is what one grasps when one understands that expression. It is this sense of ‘meaning’ which leads some to suggest that Predicate Extensionality is “not very intuitive.”² For when I grasp the meaning of ‘is blue’, for example, I don’t ordinarily think of the

¹Though, see Simmons (2000).

²See Bach (1989, 8), for example.
set or class of all blue things. Predicate Containment is similarly unintuitive. For when I judge
that Fred is blue, I don’t ordinarily mean that Fred is a member of the set or class of blue things.
In other words, logical extensionalism is not psychologically realistic—the meanings it assigns to
expressions do not closely align with our understanding of those expressions, as long as we assume
that what we seem to grasp tracks what we actually grasp. But if what we seem to grasp tracks what
we actually grasp, then our first-person perspectives as language-users can help guide us toward a
more psychologically realistic semantics. It may eventually yield philosophical insights about our understanding of
singular terms, counterfactual reasoning, and modal and mathematical cognition generally. And a
semantics which better reflects the structure of human understanding may ease the burden on logic
students learning to reason about reasoning for the first time.

What I grasp when I grasp the meaning of ‘is blue’ is either the property of being blue itself or the
concept of that property. Both concepts and properties are intensional entities. Although intensional
entities also have extensions, the sets or classes of actual things to which they apply, intensional
entities may differ even if their extensions are identical and contain all the same individuals. For
instance, consider the properties of being a cordate and being a renate. Even if we suppose they
have identical extensions, the properties themselves are distinct. Since these properties are distinct,
the concepts of them are also distinct. We grasp a difference in the meanings of the predicates ‘is a
cordate’ and ‘is a renate’, and assigning concepts or properties as the meanings of these predicates
can capture this difference. But assigning the same extension as the meaning of both predicates
dispenses over the difference.

The traditional alternative to logical extensionalism assigns intensional entities instead of exten-
sions as the meanings of predicates:

**Predicate Intensionality.** The semantic value of the predicate is an intensional entity.

Predicate Intensionality may preserve the difference in meaning for coextensional predicates such
as ‘is a renate’ and ‘is a cordate’, but it and Predicate Containment form an unstable pair of
commitments in a semantic approach. For if the meaning of a predicate isn’t its extension but instead
some intensional entity, it either isn’t the sort of thing that has members, which immediately violates

---

3 Quine (1951, 21).
Predicate Containment, or, if it does have members, doesn’t respect the motivation for Predicate Intensionality in the first place. Let me explain this second disjunct. We can adopt both Predicate Intensionality and Predicate Containment only if we identify the meaning of any predicate with an intensional entity that is also a set whose members are the possible individuals which satisfy the predicate in at least one possible world. On such an approach, the meaning of ‘is blue’, for example, is the set of blue individuals from each possible world (or a function from possible worlds to sets of blue individuals). In the next section, I argue that such complex sets or functions do not closely approximate my understanding of predicates such as ‘is blue’. If one adopts Predicate Intensionality because it better respects our understanding of the meanings of predicates, there is pressure to replace Predicate Containment with another way to link the meanings of subjects and predicates to assign truth-values to singular propositions.

The desire for a more psychologically realistic semantics itself suggests an alternative to Predicate Containment. ‘Fred’ refers to Fred, so an understanding of what the name refers to is an understanding of Fred. What, then, is an understanding of Fred? If I learn that Fred is blue, for example, I adjust my understanding of him so that it contains his being blue. If I learn that Fred isn’t red, I adjust my understanding of Fred so that it excludes his being red. A perfect understanding of Fred as he actually is contains all and only what is truly predicable of Fred: it contains his being F if and only if Fred is F. Therefore, if we want a semantics whose meanings more closely align with our understanding of the individuals to which names refer, we may identify the meaning of a name with the perfect understanding of the name’s referent. If (i) the semantic value of a name is the perfect understanding of the name’s referent, (ii) a perfect understanding of something contains all and only what is truly predicable of it, and (iii) predicables serve as the semantic values of predicates, we may replace Predicate Containment with the commitment below:

4 Lewis (1986b, 50 ff.)

5 The conjunction of Predicate Intensionality and Predicate Containment is also problematic because the meanings of predicates will not have the kinds of members on whose basis we may straightforwardly calculate the proper truth-values in a semantics for first-order logic. Suppose, on such an approach, that the set of blue individuals from each possible world serves as the meaning of ‘is blue’ and that Tom, a possible but not actual individual, is one of its members. This approach assigns a semantic value to ‘Tom’ which is a member of the semantic value of ‘is blue’. Ordinarily, this would suffice for the truth of ‘Tom is blue’. We have stipulated that Tom is not actual, however, so ‘Tom is blue’ is false. In this case, Predicate Containment and Predicate Intensionality fail to cover the falsity of ‘Tom is Blue’ on the basis of the meanings of ‘Tom’ and ‘is blue’.
Subject Containment. The semantic value of the subject contains the semantic value of the predicate.

The alternative to logical extensionalism is the conjunction of Subject Containment and Predicate Intensionality. It says that a singular proposition is true just in case the semantic value of its subject contains an intensional entity which is the semantic value of its predicate. I’ll call this alternative logical intensionalism and any variant of this alternative an intensional approach. Logical intensionalism has a venerable history, and linguists in natural language semantics now use it widely. The most well-known historical variant of logical intensionalism is Leibniz’s conceptual containment theory, which says that a singular proposition is true when the concept of the subject contains the concept of the predicate. Richard Montague is a much more recent proponent of logical intensionalism, and as Edward Zalta has noted, Leibniz’s containment theory anticipates Montague’s subject-predicate analysis of singular propositions as well as his analysis of quantified noun phrases. Whereas the containment theory says the semantic value of ‘Fred’ is a concept which contains the concepts Fred satisfies, Montague semantics says the semantic value of ‘Fred’ is a set whose members are the properties Fred has. Whereas the containment theory says the semantic value of ‘all men’ is the concept of being a man whose parts are those concepts which all possible men satisfy, Montague semantics says its semantic value is the set of properties all men have in common. Montague and Leibniz even agree that an intensional approach is preferable at least partly because it captures discourse about non-existents.

Each approach has its costs, however. In Montague semantics, the semantic value of a predicate is a property, which is itself defined as a function over possible worlds. In Section 2, I argue that the considerations which motivate an intensional approach in the first place count against this technical machinery in Montague semantics. The intensional entities in Leibniz’s containment theory are

---


7 Zalta (2000).

8 Montague (1973, 266–267) and C 53; P 20, respectively. The references to Montague follow the pagination in Montague (1974).
not defined as functions over possible worlds, which makes it a promising springboard for a novel intensional approach. Yet we must be mindful of its problems, too. I argue in Section 3 that there are two problems with the containment theory. First, it employs a controversial and, in my view, implausible conception of actuality in order to handle certain kinds of quantification. Second, it lacks a general principle for treating relational predicates.

Many believe the world behaves in ways that secure the theorems of first-order logic. Many use the formulas of first-order logic to represent aspects of the world and its inference rules to draw valid conclusions about the world. Since I share these beliefs and practices, and since I seek a more psychologically realistic semantics, I will develop an intensional approach to first-order logic. A brief detour through both Montague semantics and Leibniz’s conceptual containment theory will clarify the challenges for developing such an approach. After that detour, I’ll present and evaluate my alternative. The resulting approach lacks functions defined over possible worlds, avoids Leibniz’s concept of actuality, and has a general principle for relational predicates.

4.2 Montague Semantics

It is important to distinguish logical intensionalism from what ordinarily qualifies as a semantics for an intensional logic. An intensional logic is a formal system used to represent intensional features of language, which include failures to freely exchange coextensive expressions for one another salva veritate. Although ‘is a cordate’ and ‘is a renate’ have the same actual extension, ‘Tom believes that Fred is a cordate’ and ‘Tom believes that Fred is a renate’ may have different truth values. Hence, we cannot always substitute one predicate for the other salva veritate in an intensional logic with non-truth-functional operators such as “believes that…” . An intensional semantics is a semantics for an intensional logic.

Not every intensional semantics is an intensional approach, in my sense, and not every intensional approach is an intensional semantics. A standard possible-worlds approach to modal logic is an intensional semantics, in the usual sense, because it is a semantics for an intensional logic. But it is not an intensional approach, in my sense, because it does not involve commitments to Predicate Intensionality and Subject Containment. Conversely, the intensional approach I later develop is not an intensional semantics because it covers first-order logic, which is thoroughly extensional.
Yet it is possible for a semantics to be both. For instance, Montague translates English sentences into formulas of an intensional logic and then offers a semantics for that logic which satisfies both Predicate Intensionality and Subject Containment. Consequently, Montague semantics qualifies as both an intensional approach, in my sense, and an intensional semantics, in the usual sense. Montague semantics promises to treat intensional features of natural language, but one could cull some prominent features from Montague semantics to develop an intensional approach to first-order logic. I’ll explain momentarily why I’d rather not do this.

In Montague semantics, the semantic value of a predicate is a property, which is characterized as a function from possible worlds (and times) to functions from individuals to truth values. Even though two functions may assign the same truth values to the same individuals at the same times in the actual world, they may disagree about the values assigned to individuals in other possible worlds. Since the semantic values of predicates are intensional entities, Montague semantics satisfies Predicate Intensionality. Montague semantics also satisfies Subject Containment. The meaning of a name is its referent’s set of properties, and a singular proposition is true when the set of properties belonging to the subject’s referent contains the property which is the meaning of the predicate. The meaning of ‘Fred’, for example, is the set of properties Fred exemplifies. ‘Fred is blue’ is true when the set of Fred’s properties contains the property of being blue.

Montague handles universally quantified and existentially quantified propositions similarly. Like names, noun phrases such as ‘all men’ and ‘some man’ denote sets of properties. The noun phrase ‘all men’ denotes the set of properties all (actual) men have in common, and ‘some man’ denotes the set whose members are properties that some man or other has. So the semantic value of ‘some man’ will be a set with incompatible properties if there are at least two men in the domain who aren’t exactly alike. Whereas ‘every man is blue’ is true just in case the set of properties all men share contains the property of being blue, ‘some man is blue’ is true just in case the property of being blue is a member of the set of properties any man at all has.

A logical intensionalist with a desire for a more psychologically realistic semantics has good reason to want an alternative to Montague semantics. First, Montague semantics assigns meanings

---

9 The relevant papers in Montague include Montague (1973), Montague (1970a), and Montague (1970b). The presentation here relies heavily on Dowty et al. (1981).

to quantified noun phrases which fail to capture our understanding of those expressions. The Montagovian treatment of ‘some man’, for example, is the set which contains any property had by some man or other. To be sure, this treatment of ‘some man’ helps provide the right truth conditions for existentially quantified sentences. But given the properties of actual men, that set contains incompatible properties. If Fred is blue and Tom is red, the set contains the incompatible properties of being blue and being red. Such an incompatible set of properties does not capture my understanding of noun phrases like ‘some man’.

Second, Montague semantics often assigns the same meaning to expressions which seem to pick out different properties. Since no two functions have the same inputs and outputs, no distinct Montagovian properties hold of the same individuals at the same times in the same possible worlds. No possible individual is round and square, and no possible individual is red all over and blue all over. Therefore, the function defined as the property of being round and square is identical to the function defined as the property of being red all over and blue all over, which means that Montague semantics assigns the same meaning to both ‘is round and square’ and ‘is red (all over) and blue (all over)’. But the first predicate seems to pick out a geometric property and the second a color property. Similar remarks apply to predicates such as ‘is triangle’ and ‘is trilateral’ whose extensions are not empty in every possible world but which are the same nonetheless. These predicates arguably have different meanings, but Montague semantics assigns the same meaning to both. A logical intensionalist should prefer an approach which, at the very least, has the formal wherewithal to assign distinct meanings to necessarily coextensive predicates.

Finally, and more generally, there is a further question about whether Montague semantics employs intensional entities of the right sort. Some have noted that there is little or no connection between the meanings Montague semantics assigns to expressions and what some linguists call “psychological reality.” David Dowty, for instance, says that the meaning of a word in Montague semantics has “in principle nothing whatsoever to do with what goes on in a person’s head when he uses that word.”11 William Lycan agrees:

A linguist who sees semantics not as a branch of mathematics but as a branch of psychology understands his semantical theory of a natural language as being a (perhaps

11 Dowty (1979, 375).
the) crucial component of an explanation of speakers’ verbal behavior. (And the ordinary notion of “meaning,” for what that is worth, is certainly a psychological notion.) ... The semanticist qua psychologist owes us an account of how merely possible worlds considered as entities could figure in a mechanical or functional explanation of a living, breathing human’s flesh-and-blood actions—of how mere possibilia that are causally unconnected to a real person can play a role in moving that person’s mouth and hand. Because it seems that they could not. That is, it seems that the semanticist’s “worlds” must be a façon de parler or a metaphor of some kind, used perhaps heuristically to “index” belief states, not components of an actual psychological mechanism.  

Lycan believes that Montague semantics cannot explain speakers’ verbal behavior due to the the causal isolation of possible worlds. The infinite complexity and formal structure of the meanings defined in terms of possible worlds pose similar problems. Although I see nothing wrong with treating semantics as a branch of mathematics to calculate the right truth values—let the flowers bloom and all that—I also see great value in the importantly distinct kind of project of assigning meanings to expressions which reflect the form and experience of human understanding. For all its formal elegance, Montague semantics does not do that.

One might contend that the Montagovian function for ‘is crimson’ closely approximates an understanding of the property of being crimson because an understanding of the property is closely related to the ability to distinguish cases of crimson from cases of non-crimson, which is essentially what the Montagovian function does. Having an ability to distinguish cases of crimson from cases of non-crimson certainly would explain some behavior. But this is not the notion of understanding I have in mind. For suppose there are two infinite cognizers, Bill and Jill. Bill has the ability to pick out any possible crimson thing from any possible non-crimson thing because he has memorized the Montagovian function for being crimson. But Bill doesn’t know that crimson is a color, what crimson looks like, or, in general, what it is to be crimson. Jill, on the other hand, has never encountered the Montagovian function for being crimson. Instead, she knows precisely what it takes for something to be crimson; let’s say she knows the ranges of hue, brightness, and saturation for anything to

---

12Lycan (1984, 291). This has also been an ongoing concern in the work of Barbara H. Partee. See Section 5 for references and a discussion of her work.
qualify as crimson. Jill judges whether something is crimson on the basis of whether it satisfies her understanding of what it is to be crimson.

Seeing someone display the ability to distinguish crimson things from non-crimson things, as Bill does, is merely defeasible evidence for someone’s having Jill’s kind of understanding. But Bill’s kind of understanding isn’t sufficient for having Jill’s deeper understanding. Some of us have some inkling of what it takes for something to be crimson. Like Jill, when we encounter something which seems, by our lights, to have what it takes to be crimson, we judge that it is crimson. Generally, this kind of understanding of ‘F’ does not take the form of imagining possible instances of F but is instead constituted by some grasp of what it takes to be F, which is, at bottom, something like a list of conditions, however imperfect or implicit that list may be. If we want a semantics whose meanings more closely align with Jill’s understanding rather than Bill’s, we will assign something besides Montagovian properties as the meanings of predicates.

Unlike Montague semantics, Leibniz’s containment theory does not define its meanings in terms of extensions across possible worlds. Instead, Leibniz’s meanings are concepts which contain other concepts. This containment structure resembles the kind of structure embodied in Jill’s understanding. This feature of the containment theory makes it a promising springboard for a more contemporary intensional approach. Yet there are three problems with the containment theory. First, it requires a controversial view of actuality to account for contingently false existentially quantified propositions. Second, it requires the same view of actuality to cover contingently true universally quantified propositions. Finally, it also lacks an explicit principle for treating relational predicates. Any adequate intensional approach must avoid these problems with quantification and relations. With an eye towards avoiding these problems in my own approach, I explain them in the next section.

4.3 The Containment Theory

Leibniz’s containment theory says that a categorical proposition is true when the concept of the subject (the subject-concept) contains the concept of the predicate (the predicate-concept). In a relatively early work, Leibniz explains how he understands this conceptual containment:

\[
\text{\ldots all gold is metal; that is, the concept of metal is contained in the general concept of gold regarded in itself, so that whatever is assumed to be gold is by that very fact assumed}
\]
to be metal. This is because all the requisites of metal (such as being homogeneous to
the senses, liquid when fire is applied in a certain degree, and then not wetting things of
another genus immersed in it) are contained in the requisites of gold (C 55; P 22)

A concept contains another when the “requisites” or requirements for satisfying the contained concept are among the requisites for satisfying the containing concept. The requisites of gold include those of metal, so something satisfies the concept of being gold only if it satisfies the concept of being metal.

Tying a concept’s parts to the requisites for satisfying that concept generally means that having more parts, and hence having more restrictive satisfaction conditions, makes for a smaller extension. Conversely, concepts tend to have larger extensions with less restrictive satisfaction conditions and hence fewer parts. Leibniz writes:

For when I say *Every man is an animal* I mean that all the men are included amongst the animals; but at the same time I mean that the idea of animal is included in the idea of man. ‘Animal’ comprises more individuals than ‘man’ does, but ‘man’ comprises more ideas or more attributes: one has more instances, the other more degrees of reality; one has the greater extension, the other the greater intension.

This inversion lies at the heart of the difference between logical intensionalism and logical extensionalism. The extension of ‘animal’ subsumes and outnumbers the extension of ‘man’. So an approach which uses extensions will presumably say that ‘every man is an animal’ is true when the extension of ‘man’ is a subset or subclass of the extension of ‘animal’. Conversely, *being a man* subsumes and outstrips *being an animal*—there is more to being a man than merely being an animal. So an approach which employs intensional entities will presumably say that ‘every man is an animal’ is true when *being a man* contains *being an animal*.

Propositions such as ‘every man is an animal’ and ‘all gold is metal’ are instances of the universal affirmative, which is Leibniz’s paradigmatic categorical proposition. Even though many propositions lack the simple form of the universal affirmative, Leibniz appeals to conceptual containment to analyze each of the six types of categorical proposition (which I list below accompanied by examples):

---

13A vi, I, 190.

14NE, IV, xvii, 8, p. 486.
Consider the singular affirmative ‘Socrates is tall’. Socrates is unique, so Socrates is tall if and only if *every* Socrates is tall.\(^{15}\) Singular affirmatives are a limit case of universal affirmatives—something is predicated of everything of a kind and there is only one such thing. As a case of the universal affirmative, the singular affirmative receives the same treatment. So ‘Socrates is tall’ is true if and only if *being Socrates* contains the concept of being tall.

*Being Socrates* is an individual concept, the concept of being a particular individual.\(^{16}\) Leibniz famously holds that individual concepts are also complete, that “the individual concept of each person contains once and for all everything that will ever happen” to that person.\(^{17}\) For instance, everything truly predicatable of Socrates—that he was snub-nosed, that he had various philosophical conversations, that he drank hemlock, etc.—is contained in his individual concept.\(^{18}\) The containment theory requires complete concepts of individuals. For if there were no complete concept of Socrates, there would be something truly predicatable of him with no containment relation to account for it.

Leibniz’s treatment of singular and universal affirmative propositions suggests two possible treatments of singular and universal negative propositions. They differ on whether the subject-concept contains a negative predicate-concept or whether it merely excludes—i.e., fails to contain—the positive predicate-concept. For instance, we may say that ‘no book is female’ is true because the concept of being a book merely excludes the positive concept of being female or because the concept of being a book contains the negative concept of being non-female.\(^{19}\) Like Robert Adams (1994, 60), I find the latter treatment to be the “more natural and perspicuous intensional interpretation,” so I

\(^{15}\) A vi. I, 182–3.

\(^{16}\) This notion of an individual concept differs from more recent formulations. For Carnap (1947, 181) an individual concept is a function from state-descriptions to individual constants.

\(^{17}\) LA 12.

\(^{18}\) DM 13.

\(^{19}\) In some places, Leibniz opts for the former. See G vii. 208, C 63, and C 368; P 57. He opts for the latter in C 86.
will proceed as if this treatment of negative propositions is part of the containment theory. Similarly, we may suppose the singular negative proposition that Socrates is not rational is true just in case Socrates’ complete concept contains the concept of being not-rational.\textsuperscript{20}

Leibniz’s treatment of particular propositions diverges somewhat surprisingly from his treatment of singular propositions. For a proposition like ‘some metal is gold’ Leibniz cannot say that the concept of being metal contains the concept of being gold. For if the concept of being metal contained the concept of being gold, all metal would be gold, which is demonstrably false. Instead, Leibniz says that a particular affirmative is true when its subject-concept, with some consistent addition, contains the predicate-concept.\textsuperscript{21} The proposition ‘some metal is gold’ is true, then, just in case the subject-concept being metal, with some consistent addition, contains the concept of being gold. One candidate for consistent addition is the concept of being gold. Since the consistent concept of being gold and metal contains the concept of being gold, ‘some metal is gold’ is true.

Much like the universal and singular negative, I’ll simply assume the containment theory treats the particular negative as a case of conceptual containment of the negative predicate-concept rather than a case of conceptual exclusion of the positive predicate-concept. Hence, a particular negative is true when the subject-concept with some consistent concept jointly contain the negative predicate-concept.\textsuperscript{22} So ‘some metal is not gold’ is true when the concept of being metal, with some consistent addition, contains the concept of being not-gold.

A summary of Leibniz’s containment theory wouldn’t be complete without mentioning the concept of existing in the actual world, which I’ll hereafter refer to as the concept of being actual. For without this concept, Leibniz cannot account for contingently true universal affirmatives or contingently false particular affirmatives. I’ll consider the case of contingently false particular affirmatives first. According to Leibniz, ‘some A is B’ is true when the concept of being A with some consistent addition contains the concept of being B. Presumably, there are many pairs of consistent concepts which nothing whatsoever satisfies. Being a mountain is contained in the consistent concept of being a gold mountain, for instance, so the containment theory seems to imply that ‘some mountain is gold’ is true even though there aren’t any gold mountains. It initially appears

\textsuperscript{20}G vii, 211; P 115.
\textsuperscript{21}See, for example, C 55; P 23.
\textsuperscript{22}C 86.
that the containment theory cannot account for the falsity of ‘some mountain is gold’. For as long as conceptual consistency suffices for the truth of a particular affirmative, the containment theory will not be able to say of anything merely possible, whose concept is consistent, that there is no such thing.

Leibniz resolves this problem by appealing to the concept of being actual. God’s nature requires that he actualize the best of all possible worlds, which is ours. Concepts which represent individuals or situations of the best possible world contain the concept of being actual, whereas concepts which represent individuals or situations which don’t appear in the best possible world contain the concept of not being actual. Since concepts of actual individuals are consistent and contain the concept of being actual, the concept of being actual can be added consistently to the concept of any actual thing. For instance, if ‘some building is tall’ expresses the proposition that some actual building is tall, Leibniz can say that since the consistent concept of being a tall building already contains the concept of being actual, we can consistently add it to the concept of being a tall building. But since concepts of non-actual individuals contain the concept of not being actual, the concept of being actual cannot be added consistently to them. If ‘some mountain is gold’, for example, expresses the proposition that some actual mountain is gold, Leibniz can account for its falsity. The concept of being a gold mountain already contains the concept of not being actual, so the concept of being actual cannot be added consistently to it.

Leibniz also appeals to the concept of being actual to cover contingently true universal affirmatives. Without this special treatment, the containment theory would imply that every true universal affirmative is necessarily true, even intuitively contingent ones like ‘every renate is a cordate’. This is apparent once we consider Leibniz’s treatment of the negation of a particular negative (‘it’s not the case that some A is not-B’). Since the containment theory says a particular negative is true when the subject-concept, with some consistent addition, contains the negative predicate-concept, the negation of a particular negative is true just in case there is no concept we could add to the subject-concept so that the resulting sum consistently contains the negative predicate-concept. That is, the negation of a particular negative is true when the subject-concept and the negative predicate-concept are inconsistent. In the traditional square of opposition, the negation of a particular negative is equivalent to a universal affirmative (‘every A is B’). Since equivalent propositions merit the same semantic treatment, we now have a string of biconditionals:
(1) ‘it’s not the case that some A is not-B’ is true iff being A and being not-B are inconsistent.

(2) ‘every A is B’ is true iff being A contains being B.

(3) ‘it’s not the case that some A is not-B’ is true iff ‘every A is B’ is.

(1) and (2) are straightforward verdicts from the containment theory, and (3) falls out of the traditional square of opposition. From (1) and (3) we infer

(4) being A and being not-B are inconsistent iff ‘every A is B’ is true.

And then from (2) and (4) we infer:

(5) being A contains being B iff being A and being not-B are inconsistent.

If any two concepts are inconsistent, no possible thing could satisfy both. So if being A and being not-B are inconsistent, then, necessarily, every A is not not-B (and hence B). Given the left-to-right portion of (5), we may infer that if being A contains being B, then, necessarily, every A is B.23 So if the containment of the predicate-concept in the subject-concept were to cover the truth of any universal affirmative, the containment theory would imply that any universal affirmative would be necessarily true if true at all, even the intuitively contingent proposition that all renates are cordates.

We’ve already seen that a conceptual inconsistency results from adding the concept of being actual to any concept of something merely possible. Leibniz exploits this conceptual inconsistency to treat contingently true universal affirmatives.24 Let’s suppose that ‘every renate is a cordate’ expresses the proposition that all actual renates are cordates. Then, since it is actually the case that all renates have hearts, the concept of being a renate and not a cordate contains the concept of not being actual.25 So the concept of being a renate and not a cordate that is actual is inconsistent. Leibniz can then say the proposition that all actual renates are cordates is true when the complex concept of being a renate and not a cordate that is actual is inconsistent.26 This treatment doesn’t require that being a renate contains being a cordate, so it doesn’t imply that renates are cordates necessarily.

23 Ishiguro (1990, 47) and Lenzen (2004, 11).
Two problems remain, however. First, the containment theory seems to imply that some intuitively distinct concepts are nonetheless identical. Although Leibniz can employ the above strategy for contingently true universal affirmatives, his commitment to the thesis that the subject-concept of every true proposition contains the predicate-concept seems to commit him to another treatment of the very same propositions. Consider the true proposition that all actual renates are actual cordates as well as the true proposition that all actual renates are actual cordates. A commitment to the predicate-in-subject principle implies that the first is true if and only if being an actual renate contains being an actual cordate and that the second is true if and only if being an actual cordate contains being an actual renate.\textsuperscript{27} If, as Leibniz elsewhere claims,\textsuperscript{28} mutual containment implies conceptual identity, this treatment of universal affirmatives implies that being an actual renate and being an actual cordate are identical.\textsuperscript{29}

If being a renate and being a cordate are different concepts, we might have thought that adding being actual to both would preserve a non-identity. One way to account for this surprising identity would be to tie the identity of concepts more closely to their extensions defined over possible individuals. Because the concepts being an actual renate and being an actual cordate are indexed to the actual world, they have the same extension defined over possible individuals, an extension which consists of those actual individuals which are both renates and cordates. So being an actual renate and being an actual cordate would be identical if having the same extension guaranteed that concepts were identical. One of my original motivations for an intensional approach without possible worlds was to allow for distinct though necessarily coextensive intensional entities. To secure distinct though necessarily coextensive intensional entities, one must tie their identity conditions to something other than their extensions across possible worlds. This is the kind of view I will develop shortly.

The second, and, in my view, more serious problem is that Leibniz's conception of actuality—the one he uses to cover contingently false particular affirmatives and contingently true universal affirmatives—departs from a now widely accepted view of possibility. Imagine something which is

\begin{thebibliography}{9}
\bibitem{27} Ishiguro (1990, 194).
\bibitem{28} C 368; P 58.
\bibitem{29} Unless the containment theory reserves a special treatment for contingently true universal propositions, one will also agree with Swyower (1995, 103) that the containment theory wrongly implies the identity of coextensional concepts—not just being an actual renate and being an actual cordate but also being a renate and being a cordate. But here I’m granting that Leibniz can account for the non-identity of being a renate and being a cordate.
\end{thebibliography}
possible but non-actual, say someone who’s life resembles Obama’s except our imagined person lives in a universe where an asteroid on the other side of the universe has a slightly different trajectory. We’ll call him “Jack.” Jack’s complete concept, according to Leibniz, contains not only the concept of being the President of the United States but also the concept of being non-actual. The non-actuality of what Jack’s complete concept represents is not merely something we note about the concept; it is also built into the very concept. Generally, a concept’s parts are those concepts something must satisfy in order to satisfy that concept. So, if something were to satisfy Jack’s complete concept, it would also satisfy the concept of being non-actual. So nothing could satisfy Jack’s complete concept and also be actual. Jack is possible, but not possibly actual. The view that the possibility of something’s being actual and its being possible simpliciter are the very same thing is widely accepted, and, in my opinion, rightly so. But Leibniz does not seem to accept the equivalence.

Let me note one further deficiency in the containment theory from the perspective of modern logic. The containment theory says that a subject-predicate proposition is true when the subject-concept contains the predicate-concept. There is no trivial generalization from this treatment of subject-predicate propositions to relational propositions such as ‘A is greater than B’. In such a relational proposition, which is the subject: ‘A’, ‘B’, both, or neither? Leibniz provides no clear answer to this question in his logical works.

The lack of an intensional treatment of relational propositions is even more problematic against the backdrop of standard extensional approaches. These approaches interpret any n-ary predicate as a set of ordered n-tuples. Binary predicates are assigned sets of ordered pairs so that the semantic value of a binary predicate such as ‘is the father of’ is the set of father-son ordered pairs. Each monadic predicate is treated formally as a 1-ary relational predicate, i.e., the set of unit sets whose members are the individuals satisfying those predicates. The value of a monadic predicate such as (say) ‘is tall’ is the set of all unit sets of tall things. The treatments of both monadic and relational predicates follow from a general principle which assigns a set of ordered n-tuples to each n-ary predicate. It’s unclear whether a formal apparatus could cover relational predicates in an intensional and yet similarly systematic way. The intensional approach seems to take the monadic predicate as its

---

30 For related worries about relations in the containment theory, see Russell (1937, 12–15) and Parkinson (1965, 39–52).
central case, and the way it handles monadic predicates doesn’t lend itself to any obvious extension to relational predicates.

I originally turned to the containment theory to see how an intensional approach might work without possible worlds. But the containment theory implies the identity of some intuitively distinct concepts, seems to require a controversial and perhaps implausible conception of actuality, and lacks a systematic treatment of monadic and relational predicates. In the rest of the chapter, I develop an intensional approach which avoids these problematic aspects of the containment theory without recourse to possible worlds.

4.4 The Alternative

Logical intensionalism says that singular propositions are true when the semantic value of the subject contains an intensional entity associated with the predicate. This characterization leaves open the nature of these intensional entities as well as the sense in which they contain each other. To preserve this openness and streamline the presentation, let me introduce some terminology and offer my preferred interpretation of the formalism.

For brevity’s sake, I will call the approach’s intensional entities “properties.” This is nothing but a label: I don’t mean to endorse any particular account of properties or some disjunction of accounts. The roles these entities play in the semantics permit a wide range of metaphysical accounts of their nature. Nothing about their roles precludes their being Leibnizian concepts or Platonic universals or fictions or something else. Developing a formally adequate intensional approach to first-order logic doesn’t require maximal specificity about the underlying metaphysics—no more so than an extensional approach to first-order logic requires a specific metaphysical accounts of sets and material objects.

Logical intensionalism, by definition, requires a containment relation among intensional entities, but my approach uses a relation of property parthood instead. Since property containment is the converse of property parthood (being F is part of being G just in case being G contains being F), an approach is none the less intensional for adopting one or the other. Why bother with property parthood, then? In Chapter 1, I develop a property mereology which accounts for a way that we often talk about properties. We often speak as if properties have parts. We say that being an animal is
part of being a mammal, that being rectangular is part of being square, and so on. My conception of property parthood accounts for these kinds of locutions. Most importantly, it provides us a less onerous way to refer to the properties which a property contains. The properties contained in a property are its “parts.”

A semantics requires a language to interpret, and I will use a standard language \( L \) for first-order logic. Its primitive vocabulary includes:

- individual constants (names) \( a, b, ..., \) with or without numerical subscripts
- individual variables \( x, y, ..., \) with or without numerical subscripts
- for each \( n \) greater than 0, \( n \)-place predicates \( F, G, \) with or without numerical subscripts
- symbols \( \neg, \) and \( \supset, \) \( \forall, \) (, and )

Though I will not do so here, we can define the other connectives, including disjunction (\( \lor \)) and conjunction (\( \land \)), using \( \neg \) and \( \supset \). And we may define the existential quantifier (\( \exists \)) in the usual way in terms of \( \neg \) and \( \forall \). I will also use a number of metalinguistic variables to talk about the object language: \( \alpha \) (with or without numerical subscripts) for terms, i.e., constants and variables; \( \Pi \) (with or without subscripts) for predicates; and \( \phi \) and \( \psi \) (with or without subscripts) for well-formed formulas (wffs). The wffs are formed with the two rules below:

(i) If \( \Pi \) is an \( n \)-place predicate and \( \alpha_1, ..., \alpha_n \) are terms, then \( \Pi \alpha_1, ..., \alpha_n \) is a wff.

(ii) If \( \phi \) and \( \psi \) are wffs, and \( \alpha \) is a variable, then \( \neg \phi, (\phi \supset \psi), \) and \( \forall \alpha \phi \) are wffs.

As usual, an occurrence of a quantifier within a formula has a scope, and we can use the definition of a quantifier’s scope to define both free and bound variables. I’ll take this all for granted and will not provide those definitions here.\(^{31}\)

The semantic value of a name is not the individual it refers to but a property with a structure and function akin to a Leibnizian complete concept. Like an individual’s complete concept, an individual’s complete property has parts which correspond to the features of that individual. Obama is tall, for example, if and only if being tall is part of Obama’s complete property. For convenience, I’ll

\(^{31}\) Gamut (1991a, 96).
say the name ‘a’ “denotes” the complete property of being a, but by this I mean that the interpretation function returns the complete property as the semantic value of the name, not that the complete property is the name’s referent. We may think of this denoting relation as a two-part function which takes us from a name to its referent and then from its referent to the referent’s complete property.32

Now one such feature that individuals have is being identical to themselves, so an intensional approach to a language with identity should be expected to account for the identity of individuals. However, I will not reserve a special kind of treatment for the identity predicate as, say, a logical constant. Nor will I discuss in much detail how an intensional approach should treat an identity predicate. Yet I can say that the approach does not require the identity of indiscernibles or its denial. Let’s suppose the identity predicate is treated like any other. Then we can say, for example, that being identical to Obama is part of Obama’s complete property. In that case, the possibility of a complete property exactly like Obama’s in what it contains except for its identity property allows for someone to be qualitatively identical to Obama but not identical to him. Or, instead, one may define identity derivatively in terms of which qualitative properties are parts of each complete property. If someone cared to do so, one could develop models like this to preserve the identity of indiscernibles. But nothing about the approach in general requires the identity of indiscernibles or its denial.

The semantic values of predicates are not their extensions in this world or their extensions across possible worlds. Instead, the semantic value of a predicate is a property represented as the value of an abstraction function. The lambda-abstraction of the well-formed formula Fa, λxFx, is often thought of as representing the property of being an F. A lambda-abstraction of an atomic sentence “leaves a hole” where there was previously a name. We may abstract from a in the sentence Fa to an expression in the metalanguage which represents the property of being an F. Or we may abstract from a in the sentence Rab to an expression in the metalanguage to the property of being something R-related to b. It is also commonplace to treat an expression such as λxFx as denoting the function which spits out ‘true’ or ‘1’ when the input is in the extension of ‘F’. Although the properties represented by lambda-abstractions will serve as the semantic values of predicates, lambda-abstractions will not denote functions as they commonly do. So the properties represented by lambda-abstractions will remain unanalyzed and will not be defined set-theoretically.

32We might also think of the denoting relation in Montague semantics as a two-part function. There, the semantic value of a name is the set of properties which the name’s referent satisfies.
Lambda expressions are notoriously difficult to read. I will use a simpler formalism for my simpler purposes. In the metalanguage, we form a lambda-abstraction of a wff $\Pi \alpha_1, \ldots \alpha_n$ by replacing exactly one of the terms with a special metalinguistic variable ($x$) and enclosing the result in a pair of brackets. For example, $[Fx]$ is an abstraction of any wff $F\alpha$, and in the metalanguage it denotes the property of being an $F$. I will also assume that predicates come pre-packaged with numbers assigned to their argument places. 1 is assigned to the first argument place, 2 to the second (if there is a second place), and so on, until $n$ is assigned to the last place of an $n$-ary predicate. 33

When lambda-expressions in the metalanguage denote properties, I will often call those properties property abstractions. Since we represent a property abstraction by replacing a single term from an atomic wff, there is one such property for each argument place in the wff. Any atomic wff $\Pi \alpha_1, \ldots \alpha_n$ therefore has $n$ different property abstractions.

Below, the interpretation function $I$ takes wffs with 1-ary, 2-ary, and 3-ary predicates and spits out sets of property abstractions:

$$I(F\alpha) = \{[Fx]\}$$

$$I(Gab) = \{[Gxb], [Gax]\}$$

$$I(Habc) = \{[Hxbc], [Haxc], [Habx]\}$$

In the first case, $I$ abstracts from $a$ in the proposition $Fa$ and gives a singleton set whose member is the property of being an $F$. In the second case, $I$ abstracts from each of $a$ and $b$ in turn, and gives the set whose members are the property of being an $x$ such that $x$ is $G$-related to $b$ and the property of being an $x$ such that $a$ is $G$-related to $x$, respectively. Finally, let’s suppose that $Habc$ says that $b$ is between $a$ and $c$. The interpretation function takes $Habc$ and gives a set of three property abstractions whose members are the property of being an $x$ such that $b$ is between $x$ and $c$, the property of being an $x$ such that $x$ is between $a$ and $c$, and the property of being an $x$ such that $b$ is between $a$ and $x$.

Sets of abstractions help decide a proposition’s truth value in this intensional approach. To show how, we need to introduce the notion of an intensional model. Such a model is a pair $\langle D, I \rangle$ such that:

33This way of organizing a predicate’s argument places will be problematic if there are no asymmetric relations, as, for example, Dorr (2004) has argues.
(1) \( D \) (the domain) is a non-empty set of complete properties of individuals.

(2) \( I \) is a function (the interpretation function) obeying the following constraints:

\[
(2a) \text{ if } \alpha \text{ is a constant, then } I(\alpha) \in D.
\]

\[
(2b) \text{ If } \Pi \text{ is an } n\text{-place predicate, then } I(\Pi \alpha_1, \ldots, \alpha_n) \text{ is the set } \{[\Pi \alpha_1, \ldots, \alpha_n], \ldots, [\Pi \alpha_1, \ldots, \alpha_n] \},
\]

which contains an abstraction from each argument place in \( \Pi \alpha_1, \ldots, \alpha_n \).

The domain \( D \) houses complete properties of individuals. For my own purposes, I shall restrict the domain to complete properties of actual individuals. The interpretation function \( I \) takes each name and gives its referent’s complete property. It also takes the appearance of a predicate within a wff and provides a set of abstractions. This treatment of predicates differs significantly from the standard treatment of predicates in first-order logic. Standardly, the interpretation of a predicate is \textit{formula independent}—it receives the same interpretation, its extension, no matter which formulas it appears in. Here, however, the interpretation of a predicate is \textit{formula dependent}—each occurrence of a predicate within a formula receives its own interpretation, a set of property abstractions, and which set of property abstractions the interpretation function assigns to a predicate depends on the formula that the occurrence of the predicate appears in.

We eventually want to define truth relative to any model \( \langle D, I \rangle \). Now is a good time to provide a quick glimpse at what this definition will eventually look like. Let’s start with the simplest case of atomic propositions. Any such proposition \( \Pi \alpha_1, \ldots, \alpha_n \) is true just in case for each \( m, 1 \leq m \leq n \), if \( [\Pi \alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \alpha_n] \) is a member of \( I(\Pi \alpha_1, \ldots, \alpha_n) \), then \( [\Pi \alpha_1, \ldots, \alpha_m, \alpha_{m+1}, \alpha_n] \) is part of \( I(\alpha_m) \), the complete property which corresponds to the \( m \)th term of the proposition. Some examples will help illuminate the general idea. For simplicity, I’ll represent an individual \( \alpha \)’s complete property as ‘\( [\alpha] \).’ First, consider ‘\( Fa \)’. It is true when \( [Fx] \), the property of being an F, is part of \( [a] \). Second, consider ‘\( Habc \)’. It is true when \( [Hxbc] \) is part of \( [a] \), \( [Haxc] \) is part of \( [b] \), and \( [Habx] \) is part of \( [c] \).

Now let’s consider propositions with quantifiers. The treatment of quantification here splits the difference between Montague’s and Leibniz’s treatments, on the one hand, and standard extensional treatments, on the other. The truth of quantified statements hangs on intensional entities, as Leibniz and Montague would have it. But, like standard extensional approaches, the treatment of quantified statements depends on the entities in a domain, not on something such as “the F in general.” Suppose, for example, the domain contains the complete properties \([a], [b],\) and \([c]\). The current approach
says that ‘a is F’ is true when being F is part of [a]. Similarly, ‘b is F’ is true when being F is part of [b]. The same goes for ‘c is F’ and [c]. The complete properties of a, b, and c exhaust the domain, and being an F is part of them all. Therefore, ‘∀xFx’ (read “everything is F”) is true. Everything is such-and-such when the property of being such-and-such is part of each complete property in the domain.

This intensional treatment of quantification requires the familiar formal device of a variable assignment:

\[ g \text{ is a variable assignment for model } \langle D, I \rangle \text{ iff } g \text{ is a function that assigns to each variable some complete property in } D. \]

For a wff such as ∀xFx, a variable assignment allows us to peel off the quantifier and define its truth in terms of the open formula Fx. Under a variable assignment g, x denotes g(x), some complete property in D. If [Fx] is part of the complete property g(x), then ‘Fx’ is true under the assignment g. But we want to say that ‘Fx’ is true no matter which complete property x temporarily denotes, not just when x denotes the complete property g(x). Let a variable assignment g^α_u assign some complete property u to α (which g itself might not do) and then assign to the remaining variables every complete property that g does. Consider, then, all the assignments g^x_u which may differ with g only on what to assign to x. We can use these to define the truth of ∀xFx. This wff is true under the assignment g just in case ‘Fx’ is true under every variable assignment g^x_u. In this simple case, then, ∀xFx is true just in case [Fx] is part of every complete property in the domain.\(^{34}\)

With variable assignments now in hand, we may define the denotation of any term. Let M be a model, g be a variable assignment, and α be a term. We define \([α]_{M,g}\), the denotation of α (relative to M and the assignment g), as follows:

\[ [α]_{M,g} = \begin{cases} I(α), & \text{if } α \text{ is a constant} \\ g(α), & \text{if } α \text{ is a variable} \end{cases} \]

Relative to a model and variable assignment, the denotation of a constant is whatever complete property I assigns to that constant. The denotation of a variable (relative to a model and variable

\(^{34}\)We can also use these variable assignments to cover free variables in the usual ways.
assignment) is whatever complete property the variable assignment already assigns to the variable. So, for instance, if \( I(a) = [\text{Sally}] \) and \( g(x) = [\text{Sally}] \), then \([\text{Sally}]\) may be a constant’s denotation as well as variable’s temporary denotation.

Putting many pieces together, we may now define the valuation function, which defines truth relative to a model:

The valuation function \( V_{M,g} \) for a model \( M \) and a variable assignment \( g \) is the function that assigns to each wff either 0 or 1, given the constraints below:

(i) For any \( n \)-place predicate \( \Pi \) and any terms \( \alpha_1, \ldots, \alpha_n \), \( V_{M,g}(\Pi \alpha_1 \ldots \alpha_n) = 1 \) iff for each \( m, 1 \leq m \leq n \), \( [\Pi \alpha_1, \ldots, \alpha_{m+1}, \alpha_n] \) is part of \([\alpha_m]_{M,g}\).

(ii) For any wffs, \( \phi, \psi \), and any variable \( \alpha \):

   (iia) \( V_{M,g}(\neg \phi) = 1 \) iff \( V_{M,g}(\phi) = 0 \)

   (iib) \( V_{M,g}(\phi \supset \psi) = 1 \) iff either \( V_{M,g}(\phi) = 0 \) or \( V_{M,g}(\psi) = 1 \)

   (iic) \( V_{M,g}(\forall \alpha \phi) = 1 \) iff for every \( u \in D \), \( V_{M,g^u}(\phi) = 1 \)

I illustrate how this machinery works in Appendix 1. There, I build some intensional models and calculate the truth values of formulas with multiple quantifiers.

Let me mention a few distinctive features of intensional models. Classical first-order logic includes both the law of non-contradiction and the law of excluded middle as theorems. In extensional approaches, two features of sets help secure these theorems. First, on no model is it true that an individual both F and not-F since no set has an individual as a member and also and lacks an individual as a member. This helps secure the law of non-contradiction. Second, on no model is it true that an individual neither F nor not-F since, for any individual, the set of Fs either contains or does not contain the individual. This helps secure the law of excluded middle. In my intensional approach, two analogous features of complete properties help secure both theorems. On no model is it true that an individual both F and not-F since no complete property both has as a part and lacks as a part the property of being F. And on no model is it true that an individual is neither F nor not-F since any complete property either has the property of being F as a part or lacks the property of being F as a part.
Extensional approaches may represent certain metaphysical impossibilities (something is both red and green all over) and logical impossibilities (it is and isn’t the case that Fred is tall) within its models. But they cannot represent that David has the property of being Solomon’s father and also that Solomon lacks the property of being David’s son. For David has that property only if the ordered pair ⟨David, Solomon⟩ is a member of the set of father-son pairs, and as soon as we insert that ordered pair into the set of father-son pairs to tag David as Solomon’s father, Solomon comes along for the ride as David’s son. Every classical extensional model is harmonious—i.e., if a model says that α is Π-related to β then the same model guarantees that β is something such that α is Π-related to it.

The current approach allows for inharmonious models. Consider the proposition Gab, which is true when [Gxb] and [Gax] are parts of [a] and [b], respectively. So we should expect Gab to be false when either [Gxb] isn’t part of [a] or [Gax] isn’t part of [a]. This opens up the possibility of models in which [Gxb] is part of [a] but [Gax] isn’t part of [b]—models in which, say, being Solomon’s father is part of being David but being David’s son isn’t part of being Solomon. Now if soundness or completeness were on the agenda, we might simplify matters and restrict ourselves to the set of harmonious intensional models. These models are such that for any wff Πα₁...αₙ, [Πα₁, ... x, αₘ₊₁, αₙ] is part of [αₘ] for some m, iff for every m, 1 ≤ m ≤ n, [Πα₁, ... x, αₘ₊₁, αₙ] is part of [αₘ]. That is, harmonious intensional models are ones in which for every wff, either each abstraction from that formula is part of the appropriate complete property or no abstraction is.

The current approach also diverges from the containment theory’s treatment of negation. In Section 3, I said that ‘Fred is tall’ is true according to the containment theory when the subject-concept of being Fred contains the negative predicate-concept of being not-tall. The current approach, however, says that a negated atomic proposition such as ¬Fa is true when Fa is false. And Fa is false when Fa isn’t true, i.e., when [Fx] isn’t part of [a]. So ¬Fa is true when [Fx] isn’t part of [a], not when [a] contains the negative property of being not-F. Yet if one would like to distinguish sentence-level negation from predicate-level negation, we can easily expand the current semantics and combine the containment theory’s treatment of predicate-level negation with the current treatment for sentence-level negation.

35To be more precise, in these models [Gax] would represent the property of being an x such that David is the father of x. I’ve adopted the less precise, potentially misleading, but more concise reading here and in what follows.
4.5 Evaluation

Before I examine whether my semantics resolves those problems with the containment theory, it is important first to verify that it meets the criteria for an intensional approach. An intensional approach involves two commitments for treating singular propositions:

*Subject Containment.* The semantic value of the subject contains the semantic value of the predicate.

and

*Predicate Intensionality.* The semantic value of the predicate is an intensional entity.

My semantics does not obviously meet either of these criteria. The semantic value of the predicate ‘F’ seems to be the singleton set \{[Fx]\}. Sets aren’t intensional entities, so my approach appears to violate Predicate Intensionality. What’s more, the sets which seem to be the values of predicates are not parts of (or contained in) complete properties. So my approach apparently violates Subject Containment, too. However, this situation is similar for standard extensional approaches, which assign a set of n-tuples to each n-ary predicate. Standardly, the semantic value of each monadic predicate is not the set of individuals which satisfy the predicate, say \{Bill, John, Suzy\}, but the set of unit sets of those individuals, say \{\{Bill\}, \{John\}, \{Suzy\}\}. Since the semantic value of a name is an individual and not a unit set, the semantic value of the predicate doesn’t contain the semantic value of the subject. These standard treatments appear to violate both Predicate Containment and Predicate Extensionality similar to the way that my approach seems to violate Subject Containment and Predicate Intensionality.

Both approaches apparently violate their respective commitments due to their general principles for treating predicates. To formulate such a general principle, one must treat monadic predicates as if they were 1-ary relational predicates. Treating monadic predicates as 1-ary relational predicates is a useful formal convention. Like many textbook presentations of extensional approaches, we could easily give up that convention and treat monadic and relational predicates differently. We could assign sets of property abstractions to relational predicates, on the one hand, and assign a stand-alone property abstraction to each monadic predicate, on the other. The resulting approach
would satisfy Predicate Intensionality since the meaning of a predicate in a singular proposition would be a property. And since the meaning of a subject would contain (or not contain) the property which is the meaning of its predicate, the approach would thereby also satisfy Subject Containment. There is no substantial differences between my semantics and the one just jiggered here. The sets appearing in the values of the interpretation function are mere scaffolding, ensuring that the relations among properties determine the correct truth conditions.

My semantics is not only intensionally kosher, it also avoids the problems with Leibniz’s containment theory. First, it does not require Leibniz’s controversial conception of actuality. And, secondly, it has a well-defined principle for handling predicates of any arity. Let’s look first at how my approach accounts for contingently true universal propositions without recourse to Leibniz’s concept of actuality.

Consider once more the previously problematic renate-cordate proposition. Without a special treatment for this contingent proposition, the containment theory would imply that it is necessarily true. On the present approach, we may represent ‘all renates are cordates’ as $\forall x(Fx \supset Gx)$. The formula is true just in case every complete property in the model is such that if $[Fx]$ is part of it, so is $[Gx]$. So the approach does not imply that there is a containment relation between $[Fx]$, being a renate and $[Gx]$, being a cordate. Consequently, the approach does not imply that renates are cordates necessarily. Similar remarks apply to the formula when we switch ‘F’ and ‘G’ to represent the proposition that all cordates are renates. Hence, the renate-cordate propositions are true on the current approach because being a renate and being a cordate are parts of the same complete properties of individuals, not because being a cordate and being a renate are parts of each other (and hence identical).

Yet there are two concerns about whether my approach fares better than Leibniz on this front. Let’s grant the controversial assumption that there is a complete property for every possible individual. If these complete properties are in the domain, $\forall x(Fx \supset Gx)$ will be true if and only if $[Gx]$ is part of every one of the complete properties that $[Fx]$ is. Then it will be hard to deny that, necessarily, every F is also a G if it is true that every F is indeed a G. So if there is no way to screen out complete properties of merely possible individuals from the domain, the current approach does no better than the containment theory of accounting for contingently true universal propositions. But of course there is a way to screen out complete properties of merely possible individuals. We may restrict the
domain however we like. Extensional approaches often restrict the domain to actual individuals or some subset of them. The current intensional approach has a similar freedom. Since we may restrict the domain to complete properties of actual individuals, there’s no pressure to admit that universally quantified propositions are true only if they’re necessarily true.

The second concern is that if a complete property of an individual contains the property of being F, the individual is therefore necessarily or essentially F. If this were true, the distinction between contingently true and essentially or necessarily true singular propositions would collapse. Avoiding this collapse is an exercise in two parts. In Chapter 2, “Modal Intensionalism”, I extend the intensional approach to modal propositional logic and argue that a proposition $p$ is necessarily true if and only if being such that $p$ is part of being a world in general. Even if being tall is part of Fred’s complete property, if the propositional property of being such that Fred is tall is not part of being a world in general, then it is not necessary that Fred is tall. I see no reason to think that being such that $\text{Fred is tall}$ is part of being a world in general and no reason to think that it is necessary that Fred is tall.

Of course, being such that $\text{Fred is tall}$ may not be part of being a world in general because Fred himself is not a necessary being. If he isn’t necessary, he isn’t necessarily tall either. But Fred’s contingent existence doesn’t guarantee the contingency of his height. Fred might be a contingent being whose existence requires his being tall. In Chapter 4, “Modal Idealism,” I provide a semantics for quantified modal logic which captures these kinds of contingent truths. The main idea involves a distinction between an individual’s complete property, which captures all the actual features of the individual, and its individual property, which captures the essential features of the individual. If that semantics is viable, we can table the concern that an individual is essentially such-and-such merely because being such-and-such is part of its complete property.

The containment theory also uses the concept of being actual to cover contingently false particular affirmatives (e.g., ‘some mountain is gold’). Without that concept, the containment theory appears unable to say of anything merely possible (e.g., a golden mountain), whose concept is consistent, that there is no such thing. Although modern logicians may add an existence predicate for a variety of reasons, logicians operating within an extensional framework may simply restrict the domain to actual individuals. As a result, they may use the existential quantifier to distinguish what is from what isn’t. My approach doesn’t need Leibniz’s concept of being actual or, for that matter, anything like
it. Let ∃x(Fx ∧ Gx) represent the proposition that there is some mountain that is also gold. Insofar as we’ve already restricted the complete properties in the domain to complete properties of actual individuals, the formula will be true just in case being an F and being a G are part of some complete property of an actual individual. If there’s no such complete property of an actual individual, the formula is false—even if being an F and being a G are possibly co-exemplified.

Finally, Leibniz gives us no explicitly clear instructions for dealing with relational propositions. The current approach treats relational propositions as if each relatum were a subject: each relatum’s complete property contains a monadic property with a relational aspect—the relational aspect is the relatum’s point of view from within the relation. For instance, take the relational proposition Gabri. This proposition is true when (i) the property of being an x such that x is G-related to b is part of being a, and (ii) the property of being an x such that a is G-related to x is part of being b. Although Leibniz doesn’t give clear instructions for calculating the truth-values of relational propositions, the approach I’ve adopted here is Leibnizian in spirit. In his paper, “Primae Veritates” he writes: “There are no purely extrinsic denominations which have no foundation at all in the very thing that is denominated. For it is necessary that the notion of the subject denominated involve the notion of the predicate.”36 In other words, if there are relational truths, then those truths must somehow inhere in the notions of the things involved.

My semantics not only resolves those problems with the containment theory. From an intensional perspective, it also lacks the problematic features of Montague semantics. An intensional approach provides meanings which more closely approximate our understanding of what various expressions pick out in the world. The meanings within Montague semantics do not always closely approximate our understanding of what various expressions pick out in the world. First, Montague semantics says that an incompatible set of properties is the meaning of the noun phrase ‘some man’. But I doubt that an incompatible set of properties adequately characterizes my understanding of that noun phrase. My understanding of what the phrase purports to pick out in the world is an understanding of an individual which is also a man. My approach better captures what I take my own understanding to be, albeit under some different terminology, that there is some individual whose complete property includes the property of being a man.

Montague semantics defines a property in terms of its extension in each possible world. Therefore, no two properties have the same extension in every possible world. I don’t define properties in terms of possible worlds, so unlike Montague semantics, there’s no pressure within the view to identify necessarily co-extensive properties on the basis of their being necessarily co-extensive. Additionally, we can also flesh out the view of properties within the semantics to explain why some necessarily co-extensive properties are not identical. On my view, a property’s identity is determined by what parts it has. Even if being \( F \) and being \( G \) are necessarily coextensive, a difference in parts renders them different properties. Take, for example, the property of being a round square and the property of being red and blue all over. No possible individual has either one, which means they’re necessarily co-extensive. My view allows us to distinguish them if they have different parts. A plausible case can be made that they do have different parts. Being round is part of being a round square but not part of being red and blue all over; being red all over is part of being red and blue all over but not part of being a round square. Different parts, different properties. Or take the properties of being triangular and being trilateral, which are necessarily co-extensive. My approach permits the following explanation for their necessary coextension without appealing to their identity. All the complete properties of possible individuals which have one as a part also have the other as a part. Being triangular and being trilateral are not parts of each other but are both proper parts of the complete property of every possible triangle.

Finally, you might recall from Section 2 the cases of Bill and Jill, two infinite cognizers with different kinds of understanding of the property of being crimson. Bill has memorized the Montagovian function for ‘is crimson’, which allows him to to discriminate possible crimson things from possible non-crimson things. Jill doesn’t know the Montagovian function for being crimson; but she does know what it takes for something to be crimson. Jill judges whether something is crimson on the basis of whether it satisfies her understanding of what it is to be crimson. I sought a semantics whose meanings more closely aligned with Jill’s understanding rather than Bill’s. My semantics offers meanings which are more like Jill’s understanding of the property of being crimson.

Consider, first, a longer passage from Barbara Partee:

In earlier writings I had raised some worries about the possible psychological reality of possible worlds which in retrospect I think arose partly form misguided concern
with trying to fit possible-worlds theory “inside the speaker’s head”: but we don’t try
to fit the theory of gravitation “inside a falling apple”. Nevertheless, linguists do care
about mechanisms, and if possible-worlds semantics is a reasonable theory about the
language user’s semantic competence, there still have to exist some internal mechanisms
partly by virtue of which the external theory is correct. . . . Just to give one sample
of the difference between external theory and internal mechanism as I now see the
distinction, consider the question of how many possible worlds there might be. Even on
a non-absolutist view where we don’t take that question to have a single determinate
answer, on many applications there will be at least non-denumerably many possible
worlds. But that does not mean that a theory of internal mechanisms has to provide a way
to represent each of non-denumerably many possible worlds: for our discriminations
of possible worlds from each other are always at the level of discriminating certain
sets of possible worlds. . . . In any case, an internal mechanism might be classifying
possible worlds into sets on just some finite number of parameters at the same time that
a correct external description of the corresponding belief, assertion or whatever involved
quantification over an infinite set of possible worlds.37

The worry Partee alludes to above is that Montague semantics seems fit only for infinite minds.38
If meanings were functions defined over an infinite number of possible worlds, our finite minds would
be hard-pressed to comprehend them, let alone learn them as children. But as Partee and Dowty
explain, Montague semantics is a branch of mathematics that maps out the truth conditions of English
sentences; it isn’t meant to be part of a compelling explanation for how we comprehend or learn the
meanings of expressions from the inside. What’s more, we can accommodate our finitude to some
extent in Montague semantics by defining the meanings of English expressions as partial functions.39
Or we can think of Montagovian meanings “as representing a kind of super-competence: what we
would be like if not limited by finite brains and finite experience (e.g., if we were God.).”40

38 Hall-Partee (1979).
39 Hall-Partee (1979).
40 Hall-Partee (1979).
complexity of meanings in Montague semantics is not quite the worry that interests me.\textsuperscript{41} For the meanings in my semantics above might require an analogous kind of complexity. What interests me will be easier to express if we go ahead and assume that meanings (whether Montagovian or not) represent a God-like kind of super-competence, a full understanding of the relevant expressions. Besides, as Partee explains elsewhere, when “we make assertions we commit ourselves to their truth-conditions whether we fully ‘grasp’ them or not. We can intend more than we can master, we can aim at more than we can hit.”\textsuperscript{42} If meaning is what we grasp when we understand an expression but we also want our meanings to account for truth-conditions, then we will only partially grasp the meanings of an intensional approach. But for all that, Jill’s kind of super-competence rather than Bill’s is a more accurate portrayal of the “internal mechanisms” of our linguistic understanding.

Indeed, I want to suggest now that Jill’s kind of super-competent understanding of what predicates denote can serve as the “internal mechanism” Partee is after. To make the point, consider the determinate-determinable relation. To be a determinate of a given determinable, a property must have certain features.\textsuperscript{43} The properties of being specific kinds of triangles are determinates of the determinable property of being a triangle, and being any specific kind of triangle includes having three sides and having three angles. But a determinable’s determinates also vary with respect to each other.\textsuperscript{44} Being this or that kind of triangle includes having different side lengths and different kinds of angles.

Relative to a determinable, a full understanding of any one of its determinates includes not only a full understanding of its determinable but also whatever else is involved in being that kind of determinate. What it is to be crimson involves a little more than it does for something to be red. A full understanding of being crimson includes a full understanding of being red but it also includes a little more—being within a smaller range for hue or saturation, for example. Now I identify the semantic value of ‘is crimson’ with what I’ve called the “property” of being crimson. My intensional approach partially depends on a property mereology according to which the property of being crimson has other properties as parts, including the property of being red. Indeed, in my property mereology,

\textsuperscript{41} Partee (1989) has since give up this criticism.
\textsuperscript{42} Partee (1988, 51–52).
\textsuperscript{43} Funkhouser (2006, 551) calls these features \textit{non-determinable necessities}.
\textsuperscript{44} Funkhouser (2006, 551)
each determinable is part of every one of its determinates. So a full understanding of a determinate includes a full understanding of all that determinate’s determinables; a determinate includes all its determinables as parts. The parts of a property being \( F \) mirror a subject’s super-competent understanding of the predicate ‘\( F \)’.

Montague semantics says that the meaning of a predicate like ‘is red’ is an infinite function from possible situations to functions from individuals to truth values. That is, Montague semantics says that the meaning of ‘is crimson’ is a function which says of each possible individual whether it is crimson or not. But when we reflect on our own understanding of terms, this seems like a hollow way of thinking about meaning. An understanding of what it is for something to be crimson includes that it falls within a certain range of brightness, hue, and saturation, and also that it is red, and even colored. If a super-competent subject understood these conditions, he or she would be able to tell whether any possible individual is red. It is in virtue of a perfect understanding of these conditions, construed basically and not set-theoretically, that a super-competent subject would be able to differentiate crimson things from non-crimson things in every possible world. Given a kind of super-competent understanding of what it takes to be crimson, the correct “external” description of the meaning of ‘is crimson’ might be an infinite function from possible worlds to functions from possible individuals to truth values. But there might be only a few parameters on the basis of which a super-competent being would judge of any possible thing whether or not it is crimson—having some range of hue, having some range of brightness, and having some range saturation. These parameters, we might suppose, are parts of being crimson. If understanding a property includes understanding its parts, then the understanding of those parts may provide the “internal mechanisms” Partee seeks.

Although Montague semantics is a highly successful formal theory for mapping out the truth conditions for sentences of natural language, my semantics offers a more respectable story about our understanding of certain kinds of expressions from the inside, especially ones that denote determinables and determinates. Of course, we probably hardly ever fully comprehend the meanings of terms—we aren’t super-competent. In my own case, my phenomenology of understanding a predicate often includes nothing more than a basic—albeit imperfect—understanding of a few conditions. My semantics is not a replacement for Montague semantics. But it complements Montague semantics well. Montague semantics provides what Partee calls an “external” characterization of meanings, a function which simply says “yes” or “no” about whether an object satisfies the predicate. But
Partee also seems to hold out hope for a semantics which provides an “internal” characterization of those meanings, something that more closely resembles the underlying psychological reality. We have two more or less complementary pictures about what meanings are. If I’m right about this, perhaps linguists might find inspiration for an alternative approach within the semantics I’ve offered for first-order logic.

The intensional approach developed here meets all the criteria I sought. Like Leibniz’s containment theory, it avoids Montague’s possible-worlds account of intensional entities. But unlike Leibniz’s approach, it contains a general principle for treating predicates (both monadic and relational) and doesn’t require a controversial concept of actuality.

4.6 Conclusion

I’ve offered a variant of logical intensionalism without possible worlds that overcomes some of the problematic aspects of Leibniz’s containment theory. An intensional approach like mine is meant to assign meanings to expressions which closely approximate our understanding of what those expressions pick out in the world. Often, we understand expressions even though they don’t seem to pick out anything in the world. For instance, we often discuss what does not exist. Montague and Leibniz each tout the ability of their semantics to account for discourse about non-existents. My approach can also be extended to make sense of some discourse about non-existents. I explain in Appendix 2 one way to do this.

Non-existents also figure importantly in discourse about possibilia, especially mere possibilia. Linguists and philosophers have used the apparatus of possible-worlds semantics to capture this discourse. Yet possible worlds semantics is not an ideal approach to modal logic if we want meanings which closely approximate our understanding of various modal claims. According to possible-worlds semantics, a proposition is necessary when it’s true in every possible world. But we do not judge that something is necessary on the basis of observing its truth in every possible world. The phenomenology of modal judgment and our cognitive limitations bear this out. In the next chapter, I combine my semantic approaches for modal propositional logic and first-order logic to provide an intensional semantics for quantified modal logic.
4.7 Appendix 1: Examples

Let’s run through a few examples. It will be especially beneficial to see how the semantics handles relational predicates within the scope of multiple quantifiers.\(^{45}\)

**Example 1.** \(\exists x \exists y R_{xy}\) (“someone loves someone”)

Let’s suppose that our model has the following features:

\[ D = \{[a], [b], [c]\} \]

[a]’s parts include \([Rx_b]\).

[b]’s parts include \([Rax]\).

Now let’s suppose that \(g(x) = [a]\) and \(g(y) = [b]\). Since \([Rx_b]\) is part of \([a]\) and \([Rax]\) is part of \([b]\), \(V_{M,g}(Rx_y) = 1\). Since \(g = g[^y]_{[b]}\), \(V_{M,g[^y]_{[b]}}(Rx_y) = 1\). Therefore, there is some \(u \in D\) such that \(V_{M,g[^y]_{[b]}}(Rx_y) = 1\). So \(V_{M,g}(\exists y Rx_y) = 1\). Also, since \(g = g[^x]_{[a]}\), \(V_{M,g[^x]_{[a]}}(\exists y Rx_y) = 1\). Therefore, there is a \(u \in D\) such that \(V_{M,g[^x]_{[a]}}(\exists y Rx_y) = 1\). Hence, \(V_{M,g}(\exists x \exists y Rx_y) = 1\).

**Example 2.** \(\exists x \forall y R_{xy}\) (“someone loves everyone”)

Suppose there is a model such that:

\[ D = \{[a], [b], [c]\} \]

[a]’s parts include \([Rx_a], [Rx_b], [Rx_c]\), and \([Rax]\).

[b]’s parts include \([Rax]\).

[c]’s parts include \([Rax]\).

\(^{45}\)More illuminating examples will be possible once we introduce the existential quantifier. Since \(\neg \forall \neg \phi\) and \(\exists \phi\) are logically equivalent, we may now use (iia) and (iii) to lay out the truth conditions for the derivative existential quantifier. First, \(V_{M,g}(\neg \forall \neg \phi) = 1\) iff \(V_{M,g}(\forall \neg \phi) = 0\). Second, \(V_{M,g}(\forall \neg \phi) = 0\) iff it isn’t the case that for every \(u \in D\), \(V_{M,g[^u]}(\neg \phi) = 1\). It isn’t the case that for every \(u \in D\), \(V_{M,g[^u]}(\neg \phi) = 1\) iff there is some \(u \in D\), \(V_{M,g[^u]}(\phi) = 1\). Hence, \(V_{M,g}(\neg \forall \neg \phi) = V_{M,g}(\exists \phi) = 1\) iff there is some \(u \in D\), \(V_{M,g[^u]}(\phi) = 1\). So if \(\exists x F x\) says that (say) something is tall, then it is true when the property of being tall is part of some individual’s complete property in the domain.

78
Now for any assignment $g$ such that $g(x) = [a]$:  

$$V_{\mathcal{M},g_{[a]}^y}(Rxy) = 1 \text{ because } [Rx] \text{ is part of } [a] \text{ and } [Rax] \text{ is part of } [a].$$

$$V_{\mathcal{M},g_{[b]}^y}(Rxy) = 1 \text{ because } [Rx] \text{ is part of } [a] \text{ and } [Rax] \text{ is part of } [b].$$

$$V_{\mathcal{M},g_{[c]}^y}(Rxy) = 1 \text{ because } [Rx] \text{ is part of } [a] \text{ and } [Rax] \text{ is part of } [c].$$

So if $g(x) = [a]$, every $u \in \mathcal{D}$ is such that $V_{\mathcal{M},g_{[a]}^y}(Rxu) = 1$. Therefore, $V_{\mathcal{M},g}(\forall y Ryx) = 1$. Now if $V_{\mathcal{M},g}(\forall y Ryx) = 1$ given that $g(x) = [a]$, there is some $u \in \mathcal{D}$ such that $V_{\mathcal{M},g_{[a]}^u}((\exists y Ryx) = 1$. Therefore, $V_{\mathcal{M},g}(\exists x \forall y Ryx) = 1$.

**Example 3.** $\forall x \exists y Ryx$ (“everyone loves someone”)

Our model has the following features:

$$\mathcal{D} = \{[a], [b], [c]\}$$

- $[a]$’s parts include $[Rx]b$ and $[Rc]x$.
- $[b]$’s parts include: $[Rx]c$ and $[Rax]$.
- $[c]$’s parts include: $[Rx]a$ and $[Rbx]$.

Now for an arbitrary assignment $g$:

- if $g(x) = [a]$, then $V_{\mathcal{M},g_{[a]}^y}(Rxy) = 1$ because $[Rx]b$ is part of $[a]$ and $[Rax]$ is part of $[b]$.
- if $g(x) = [b]$, then $V_{\mathcal{M},g_{[b]}^y}(Rxy) = 1$ because $[Rx]c$ is part of $[b]$ and $[Rbx]$ is part of $[c]$.
- if $g(x) = [c]$, then $V_{\mathcal{M},g_{[c]}^y}(Rxy) = 1$ because $[Rx]a$ is part of $[c]$ and $[Rcx]$ is part of $[a]$.

For an arbitrary assignment $g$, there is a $u \in \mathcal{D}$ such that $V_{\mathcal{M},g_{[a]}^u}(Rxu) = 1$, so $V_{\mathcal{M},g}(\exists y Ryx) = 1$. Since $V_{\mathcal{M},g}(\exists y Ryx) = 1$ for an arbitrary $g$, $V_{\mathcal{M},g_{[a]}^u}(\exists y Ryx) = 1$ for every $u \in \mathcal{D}$. Thus, $V_{\mathcal{M},g}(\forall x \exists y Ryx) = 1$.  

79
4.8 Appendix 2: Non-existents

Both Montague and Leibniz prefer an intensional approach in part because intensional approaches are better-suited for discourse about non-existents. For if one assigns existential import to the quantifiers in an extensional approach, one will also construct the domain so that it consists of actual or existent individuals only. With such a domain, there’s no straightforward way to use a name to express truly that its referent is such-and-such (e.g., ‘Santa Claus is a man’) or, less controversially, to truly express that it doesn’t even exist (e.g., ‘Pegasus doesn’t exist’). If the domain contains actual individuals only, it will not include Santa Claus and Pegasus. And if Santa Claus and Pegasus are not in the domain, the model will not account for propositions with names for them. Prima facie, an intensional approach overcomes this problem—the concept or property of being Santa Claus may exist even if Santa Claus doesn’t. We can appeal to that concept or property to account for true propositions about Santa Claus in the same way the present approach appeals to complete properties of actual individuals to account for true propositions about them: even though Santa doesn’t exist, being Santa Claus contains or doesn’t contain certain other concepts/properties.

There are many ways to make sense of discourse about non-existents. Leibniz’s own approach involves adding the concept of being actual which is contained in the complete concepts of actual individuals and excluded from the concepts of merely possible individuals. His strategy for making sense of discourse about non-existents is off limits to us if we think the complete concept of an individual is a concept such that if anything were to satisfy the complete concept, it would also satisfy the concepts which the complete concept contains. For instance, if some complete concept contains the concept of being a cyclist, then if someone were to satisfy that complete concept, he/she would be a cyclist. If a complete concept fails to contain the concept of being actual and therefore contains the concept of being non-actual, anything which would satisfy the complete concept would therefore be non-actual. So if a merely possible being (which contains the concept of being non-actual) were actual, it would be non-actual. Leibniz’s treatment of ‘is actual’ as a predicate like any other predicate—contained or excluded from the subject-concept—leads to trouble. Therefore, if an intensional approach aims to account for discourse about non-existents, it won’t treat ‘is actual’ or ‘exists’ as an ordinary predicate.
I’ll briefly present one extension to my extensional approach which seems to handle discourse about non-existents without treating ‘exists’ (or ‘is actual) in the Leibnizian way. First, we single out a very special monadic predicate ‘E’ and think of it as the existence predicate so that ‘Ea’, for instance, reads as ‘a exists”. Then we introduce both an inner and an outer domain. We redefine \( D \) as a set of properties and introduce one of its subsets \( E \) as the set of complete properties of existent individuals. \( D \) is the outer domain, and some its members will be used to account for discourse about non-existents. The inner domain \( \mathcal{E} \) will consist only of complete properties of existent individuals.

In my intensional approach, each predicate other than E is interpreted as a set of abstractions. But E will not behave like that. Instead, ‘Ea’ is true when the value of \( a \), the complete property \([a]\), is a member of \( \mathcal{E} \), the set of complete properties of existents. So unlike other predicates, the treatment of ‘E’ resembles the treatment of a predicate in an extensional approach. ‘E’, in effect, separates the complete properties of existent individuals from properties of non-actual individuals in the more inclusive outer domain.

This will also be a positive free logic in which non-existent objects have positive properties. Where \( a \) is Santa Claus and \( F \) is the the property of being jolly, we will be able to say both that Santa Claus is jolly and that he doesn’t exist: \( \neg Ea \) and \( Fa \). Both formulas will be true in a model when (i) the property of being Santa is a member in the outer but not the inner domain, and (ii) the property of being jolly is part of the property of being Santa.

Given these inner and outer domains, we may also define both inner and outer quantifiers. The outer quantifiers range over properties in the outer domain, and the inner quantifiers range over complete properties of individuals from the inner domain. We define the inner quantifiers (\( \exists^E \alpha \) and \( \forall^E \alpha \)) in terms of the outer quantifiers:

\[
\exists^E \alpha \phi \iff \exists \alpha (E\alpha \land \phi)
\]

\[
\forall^E \alpha \phi \iff \forall \alpha (E\alpha \supset \phi)
\]

If being Santa Claus, for example, is a member of the outer and not the inner domain, we can rig the model so that we can say true things about him. If being a man is part of being Santa Claus, for example, ‘Santa claus is a man’ will be true. Also, as long as being Santa Claus is a member of the outer domain, the outer existential quantifier permits the inference from ‘Santa Claus is a man’ to
the (outer) existentially quantified ‘someone is a man’, which carries no existential import. In sum, this positive free logic and the accompanying intensional semantics allows us to speak truly about non-existents as non-existents and as having various positive features. It also allows us to quantify over them without implying their existence.

Since some accounts of properties permit them to exist even though they never punch their tickets as properties of something, we can use them to make sense of discourse about about non-existents. Otherwise, there is a non-trivial difficulty with making sense of discourse about non-existent objects, especially if individuals are assigned to names as we see in extensional approaches. The current conception of “properties” also allows for both inconsistent and incomplete properties. So although one would have to adopt a non-classical logic, the intensional approach could be revised to account for propositions such as ‘the round square is a square’ as well as ones like ‘it’s not the case that either Santa has high blood pressure or that he doesn’t have it’.
CHAPTER 5

Modal Idealism

5.1 Introduction

Our experience of actual objects such as cars and dogs provides the ingredients for knowledge about those objects. I know that Fido is brown at least in part because I can see Fido and his being brown. We also seem to know that some truths are necessary, such as the truth that 2 and 2 equals 4. Plausibly, however, we don’t experience anything necessary as necessary. So if we know that something must be such-and-such or that, necessarily, all Fs are Gs, A.J. Ayer worries that,

We shall be obliged to admit that there are some truths about the world which we can know independently of experience; that there are some properties which we can ascribe to all objects, even though we cannot conceivably observe that all objects have them. And we shall have to accept it as a mysterious inexplicable fact that our thought has this power to reveal to us authoritatively the nature of objects which we have never observed.¹

We need not share Ayer’s logical positivism to observe that knowledge of what’s necessary would seemingly result not from perception but from thought—or as I’ll call that faculty—the understanding. We understand that \( p \) and infer that, necessarily, \( p \). How could we acquire knowledge that a proposition is necessary on the basis of understanding it? How is knowledge of what’s necessary possible?

Two kinds of propositions seem to wear their necessity on their sleeves, and it is less mysterious how we might know that a proposition of either kind is necessarily true. First, if contradictions cannot be true, their negations must be true. So we may infer that the negation of a contradiction is

---

¹ Ayer (1946, 73).
necessary. Second, if a proposition is analytic in the sense that it is true in virtue of meaning alone, there is no way for the world to falsify the proposition as long as the proposition means what it does. Since such propositions couldn’t have been false, we may infer that a proposition is necessary from its analyticity. Let’s grant that we sometimes acquire modal knowledge on the basis of understanding merely by grasping that a proposition is analytic or that it is the negation of a contradiction. Still, we seem to know that some propositions are necessary though they fit neither of these categories.

Consider the propositions below, each of which is neither clearly analytic nor clearly the negation of a contradiction:

(1) God exists.
(2) The number 1 exists.
(3) Something exists.
(4) $2 + 2 = 4$.
(5) Nothing is simultaneously red all over and blue all over.
(6) If something is red, it has a spatial property.

One could make a case that each proposition is either analytic, the negation of a contradiction, or not necessary to begin with. There are well known strategies for denying the truth and hence the necessity of both (1) and (2). There is a tradition of defining analyticity as derivability from a set of logical truths. On this Fregean notion of analyticity, (3) and (4) are arguably analytic. (3) is arguably analytic in this sense because it is a theorem of widely adopted systems of first-order logic. (4) is arguably analytic in this sense because it follows from theorems of second-order logic in conjunction with Hume’s Principle (the principle that the number of Fs is equal to the number of Gs if and only if there is a one-to-one correspondence between the Fs and the Gs). If we derive (3) and (4) from a set of logical truths which are individually necessarily true, then we may infer that (3) and (4) are also necessarily true. Finally, perhaps there are hidden contradictions in the negations of (5) and (6). We may then infer that (5) and (6) are necessarily true.

Even if (1) and (2) are false, many of us judge that one or the other is necessarily true. What about our understandings of (1) and (2) tempts us into judging that one or the other is necessarily true.

\footnote{Frege (1884, 88).}
true? And whether (3) or (4) is analytic in Frege’s sense does not help us resolve the puzzle of how we might know that (3) or (4) is necessarily true.\(^3\) Most of those who judge that (3) or (4) is necessarily true, do so in a state of blissful ignorance about their Fregean analyticity. What about our understandings of (3) and (4) tempts us into judging that one or the other is necessarily true? Finally, even if the negations of (5) and (6) are contradictions, many of us thoughtfully judge that (5) and (6) are necessary without seeming to derive a contradiction from their negations.

We might have thought that we could boil our apparent knowledge of what’s necessary into two groups: knowledge that something is necessarily true on the basis of its analyticity and knowledge that something is necessarily true on the basis of its being the negation of a contradiction. Since we are able to judge that some propositions are analytic or negated contractions, we might then have hoped to explain how it is possible for us know that some propositions are necessary merely on the basis of understanding them. But this strategy’s prospects look bleak. Some propositions are neither analytic in the traditional sense, nor known to be negated contradictions, nor known to be analytic in the Fregean sense. Yet we still seem to know, or perhaps justifiably believe, that they are necessarily true. When we do not rely on analyticity or contradiction, how, on the basis of understanding, do we know—or at least believe justifiably—that a proposition is necessary?

We might gain traction on this question if we explain what is understood when we understand a necessary truth and also what is understood when we understand that the truth is necessary. On one plausible interpretation of what a semantics does, a semantics for modal logic might provide both kinds of explanation. A semantics for some discourse tells us what expressions in that discourse mean. We then might expect an expression’s meaning to be what one grasps when one understands that expression. Since a semantics for quantified modal logic provides the meanings of necessary truths as well as the meaning of ‘it is necessary that…’, we might expect a semantics for quantified modal logic to tell us what one would understand when one fully understands that a necessary truth is necessary. From there, we might be able to explain how, on the basis of understanding a proposition as well as an understanding of ‘it is necessary that…’, we might know that a proposition is necessary.

\(^3\)Furthermore, (3) would only count as a logical truth if (i) we tie existential quantification to existence, a metaphysical and not purely logical concept, and (ii) we assume that since the least burdensome logical systems require non-empty domains, reality must follow suit and provide those systems with at least one thing for them to represent. Thanks to Bob Adams for this point.
Given the success of possible worlds semantics, philosophers now like to express modal judgments in terms of possible worlds and define necessity as truth in all possible worlds, i.e., all possible ways the world could have been. There are two reasons I doubt that possible worlds semantics can help explain how we might know that a proposition is necessarily true on the basis of understanding. First, there are infinitely many possible worlds. Finite creatures like us cannot judge that something is necessary purely on the basis of grasping what is true in infinitely many possible worlds alone. We have neither the cognitive horsepower to grasp them all at once nor the time to grasp them in succession. Knowledge that something is true in all possible worlds is, for finite creatures like us, likely inferred from already knowing that it is necessarily true. Kripke writes:

In the actual development of our thought, judgments involving directly expressed modal locutions (‘it might have been the case that’) certainly come earlier. The notion of a ‘possible world’, though it has its roots in various ordinary ideas of ways the world might have been, comes at a much greater, and subsequent, level of abstraction.\(^4\)

If the knowledge that some proposition is true in all possible worlds is inferred, it is either inferred from the prior knowledge that it is necessarily true, per Kripke, or from knowing something about the proposition that secures either its necessity or its truth in all possible worlds or both.

Secondly, even if we had sufficient time or cognitive horsepower to grasp an infinite number of possible worlds, influential accounts of the nature of possible worlds would, if true, severely limit our access to them. Broadly speaking, there are two influential kinds of realism about possible worlds. First, David Lewis’s modal realism identifies possible worlds with island universes—universes like ours that are spatio-temporally and causally disconnected from us and from each other.\(^5\) Second, many varieties of ersatzism associate possible worlds with abstract objects which represent ways reality might have been.\(^6\) Abstract objects are typically thought of as acausal. In either case, it’s difficult to see how we could know anything about a realm of objects from which we are causally disconnected.

\(^4\)Kripke (1980, 19, n. 18).

\(^5\) Lewis (1986b).

\(^6\) See, for example, Adams (1974), Plantinga (1974), Stalnaker (1976), and van Inwagen (1986), and perhaps Kripke (1980), too, though this is less clear to me. However, I don’t mean to suggest that each of these authors would identify possible worlds with mind-independent Platonic forms.
Although possible worlds semantics does many things well, there’s room for an alternative semantics whose meanings meet two conditions. First, there’s room for a modal semantics whose meanings are mathematically simple enough that human beings could know that a proposition is necessarily true without first knowing that the proposition is analytic or the negation of a contradiction. Second, there’s room for a semantics whose meanings on some intended interpretation are the sort of objects to which we have cognitive access. The account of necessity within such a semantics would better explain how it is possible for creatures like us to know some necessary truths on the basis of understanding—or at least make some principled judgments about what’s necessarily true—, especially in cases where we don’t seem to rely on the shortcuts of analyticity and contradiction.

In this chapter, I develop such a semantics for quantified modal logic without possible worlds, accessibility relations, or formally similar stand-ins for either one. The semantics revolves around two ideas. First, there’s the idea of being this world, the actual one we live in. A proposition \( \phi \) is true when the idea of being such that \( \phi \) is part of the idea of our world, \( \alpha \). Secondly, we can abstract away from the idea of \( \alpha \) to the idea of being a world in general. A proposition \( \phi \) is necessarily true when the idea of being such that \( \phi \) is part of the idea of being a world in general.

I develop the semantics within the historically prominent view that modal truths are ideas in an infinite mind, a mind which shares some of these ideas with us. I will argue that, compared to possible world semantics, this semantics better explains how we might acquire modal knowledge on the basis of understanding—or at least make thoughtful modal judgments on the basis of understanding.

5.2 Ideas

In ordinary cases of idea possession, we may distinguish three things. When I think of Obama, for instance, there is, first, the token mental act. Somehow or other, at some place and time, I point my mind’s eye towards Obama. On the other end of my thought is the intentional object, the thing my thought is about. In this case, the object of my thought is Obama, the man. Finally, the mental act is directed towards Obama via the intentional or representational content of the mental act. This content is what I’m calling an idea of Obama. Let ideas be the intentional or representational contents by which mental acts are directed towards their intentional objects.
Ideas seem particularly well-suited to account for knowledge of modal reality. They have representational or intentional content which determines the idea’s satisfaction conditions. Someone satisfies the idea of being a Dutch 19th century painter, for example, if and only if he or she is Dutch, lives in the 19th century, and paints. And that would be true even if van Gogh and the other Dutch 19th century painters had never existed. Hence, ideas have content in virtue of which they can represent or be about what doesn’t exist. The content of an idea which nothing satisfies nevertheless represents what something would be like if it satisfied that idea. For instance, take the idea of a golden mountain. Presumably, no such thing exists, but something would satisfy the idea if it were both gold and a mountain.

Although the merely possible does not exist, we may know about mere possibilia if we have access to ideas of things which happen to be merely possible. Some ideas represent what doesn’t exist, and those ideas represent what those objects would be like if they had existed. So long as we have access to some of those ideas, we have access to what non-actual possibilia would be like if they had existed. And we do have access to them. Ideas are the currency of thought and ultimately mind-dependent entities, not mind-independent Platonic forms. We have an intimate connection to them. With the possible exception of qualia, there is nothing else with which we have a more intimate connection. Common sense says that we not only have access to ideas but that we have access to ideas of things which happen not to exist but could have. We often reason counterfactually about what could have been or about what might happen if we do this or that. In the latter case, our ability to predict fairly reliably the results of our immediate actions suggests that we grasp some ideas of possibilia fairly well.

What is disputable, however, is whether any particular metaphysical account of ideas can explain our access to them and also how those ideas fulfill the various roles we might require of them. Given that ideas are windows into what could have been, what metaphysical theory or theories of ideas would explain our apparent ability to peer into modal reality? One promising candidate says that ideas have a structure resembling Leibnizian concepts, that they have the causal efficacy of

---

7 An idea is the intentional or representational content of a mental act and also has intentional or representational content. I’ll explain how this is possible shortly.

8 According to Husserl (1983, 309–311) the purported object of an idea or content (in a mental act) need not exist. But even so, “under content we understand the ‘meaning’ of which we say that in it or through it consciousness refers to an object of its own.”
Malebranchean ideas, and that, like both Leibnizian concepts and Malebranchean ideas, they reside in an infinite mind who then provides those same ideas as the materials of our own thoughts, though we grasp them imperfectly. I will flesh out such a theory in the context of providing an original semantics for quantified modal logic. Unlike Leibniz or Malebranche, however, I need not associate the infinite mind with the Christian God or, for that matter, the God of any of Abrahamic religion. I will only assume that there is a necessarily existent mind responsible for the being of all ideas and whose cognitive powers are infinite. I won’t assume that this mind fiddles with the natural world or cares to interfere in our daily lives. I will often speak of “God’s mind” or “divine ideas” but these descriptions do not carry any religious weight.

Ideas, on this theory, have a structure which closely resembles the structure of concepts in Leibniz’s conceptual containment theory of truth. So it will be beneficial to summarize the logical structure of Leibnizian concepts. According to Leibniz, a categorical proposition is true when the concept of the subject contains the concept of the predicate:

\[ \text{...all gold is metal; that is, the concept of metal is contained in the general concept of gold regarded in itself, so that whatever is assumed to be gold is by that very fact assumed to be metal. This is because all the requisites of metal (such as being homogeneous to the senses, liquid when fire is applied in a certain degree, and then not wetting things of another genus immersed in it) are contained in the requisites of gold (C 55; P 22)} \]

A concept contains another when the requirements for satisfying the contained concept are among the requirements for satisfying the containing concept. The requirements for anything to be gold include the requirements for anything’s being metal, so \textit{being gold} contains \textit{being metal}. Or, as we might say in modern English, “being metal is part of being gold.” Tying a concept’s parts to the requirements for satisfying that concept generally means that having more parts, and hence having more restrictive satisfaction conditions, makes for a smaller extension. Conversely, concepts tend to have larger extensions with less restrictive satisfaction conditions and hence fewer parts. Here is more, from Leibniz:

---

9 This claim of necessary existence will itself be subject to the definition of necessity I provide later in the paper.

10 A vi, 1,190, par. 80.
For when I say *Every man is an animal* I mean that all the men are included amongst the animals; but at the same time I mean that the idea of animal is included in the idea of man. ‘Animal’ comprises more individuals than ‘man’ does, but ‘man’ comprises more ideas or more attributes: one has more instances, the other more degrees of reality; one has the greater extension, the other the greater intension.\(^{11}\)

Sometimes, Leibniz discusses his theory using the notion of conceptual containment; sometimes, (like in the passage immediately above), he appeals to idea inclusion; and sometimes, he uses the notion of conceptual parthood. I will split the difference and pitch my view in terms of ideas and their parts. But I’ll use the containment or inclusion metaphors when they pay their way in added convenience. We can extrapolate from these passages the kinds of judgments Leibniz would make about which ideas are parts of which others. The idea of being red is part of the idea of being crimson, the idea of being mammalian is part of the idea of being a human, and, generally, ideas of determinables are parts of ideas of determinates.

These judgments conflict with more recent accounts of intensional entities which construe them in terms of their extensions across possible worlds. On these set-theoretical accounts, *being red* is defined either as a set or function in terms of its extension in each possible world. The set of possible crimson things is a subset of the set of possible red things, so insofar as the set-theoretical views allow for a part-whole relationship between *being crimson* and *being red*, they can capture the judgment that *being crimson* is part of *being red* but not the converse judgment that *being red* is part of *being crimson*, which I prefer.

Ideas have parts, and I follow Leibniz in defining ideational parthood in terms of mental conjunction.\(^{12}\) Ideas with parts are mental conjunctions of simpler ideas. The idea of being tall and the idea of being red conjoin to form the mental conjunction of being tall and red. The instance of ‘and’ appearing between the predicates ‘tall’ and ‘red’ signifies the phenomenon of mental conjunction. Something satisfies a mental conjunction when it satisfies each of its conjuncts. Therefore, something satisfies an idea just in case it satisfies every part of the idea. Something satisfies the idea of being red and tall, for instance, just in case it is both red and tall. Consequently, ideas with more conjuncts

\(^{11}\) Leibniz (1996, IV, xvii, Sect 8, 486)

\(^{12}\) G vii 239/P 135.
typically have smaller extensions since there are more ways something may fail to satisfy the whole idea. Ideas with fewer conjuncts typically have larger extensions since there are fewer ways to fail to satisfy the whole idea.

To differentiate mental conjunction from logical conjunction, I will symbolize the relation of mental conjunction with ‘+’. An idea is part of an idea when the first is either identical with the second or with some or perhaps just one of the second’s conjuncts. That is,

(IP) An idea $I_x$ is part of idea $I_y$ just in case

(i) $I_x = I_y$ or

(ii) $I_y = I_1 + I_2 + \ldots + I_k + \ldots$ and $I_x = I_1$ or $I_2$ or $I_k$ or $\ldots$. or

(iii) $I_x = I_n + \ldots + I_k + \ldots$, and each of $I_n$, $\ldots$, $I_k$ and $\ldots$ are conjuncts of $I_y$.

There are a few remarks I want to make about (IP). First, when an idea qualifies as part of an idea via condition (ii) or (iii), I’ll often call the first idea one of the second’s conjuncts. Sometimes, then, when an idea $I_x$ is a conjunction whose conjuncts are also conjuncts of an idea $I_y$, I’ll call $I_x$ one of $I_y$’s conjuncts.

Second, ideational parthood is not the only way ideas relate to one another. (IP) isn’t meant to capture the other kinds of relations some ideas bear to one another, even if we may use ‘part’ or its cognates to describe those relations. A couple examples should suffice. Mental disjunctions of the form being an $F$ or $G$ do not have their disjuncts as parts on my preferred sense of ‘part’. If being an $F$ and being a $G$ were parts of the disjunction of being an $F$ or $G$, then in satisfying the disjunctive idea, one would also satisfy both being an $F$ and being a $G$. Obviously, something may satisfy the disjunctive idea without satisfying both disjuncts. The ideas involved in an idea of a relation are not all parts of the idea of that relation either. For instance, consider the case of the idea of being north of the equator. The idea of being north of the equator does not have being located at 0 degrees latitude as a part, even though it is part of being the equator, which is itself somehow involved in the idea of being north of the equator. (IP) carves out only one way among many that ideas relate to one another.

Second, there is an important difference between mental conjunction and logical conjunction. Although there are logical conjunctions of a proposition and itself, there are no mental conjunctions of an idea and itself. To put the same point in terms of an idea’s parts: no idea has a part twice over.
Take, for example, a purported case of repetition, say *being red and being red*. The repetition adds nothing by way of intentional or representational content. We may construe this purported case of an idea with a part twice over instead as two mental acts, two thoughts, whose content—the idea of being red—is the same in each.

Or consider a mental conjunction such as the idea of being crimson and scarlet, an idea nothing could possibly satisfy but an idea, nonetheless. The idea of being red (let’s say), is part of both. But we need not construe the idea of being crimson and scarlet as having the idea of being red as a part twice over. We may instead construe this purported repetition in one of two ways. We may borrow the previous strategy and construe it as the appearance of an idea in two mental acts, a thought of the idea *being crimson* and a separate thought of the idea *being scarlet*. Or we may construe the purported repetition of the idea of being red as how many times it is a part of salient ideas which are themselves parts of *being crimson and scarlet*. Picking out the idea of being crimson and scarlet with ‘being crimson and scarlet’ makes the ideas of being crimson and being scarlet salient parts of that idea. And the idea of being red is part of both. This overlap makes it seem on the surface that the idea of being red is a part twice over. But it is not. The idea of being crimson and scarlet is a mental conjunction whose conjuncts include the idea of being red, plus whatever other parts the idea of being crimson has, plus whatever other parts the idea of being scarlet has. Nothing could possibly exemplify the idea, but there is no repetition. Mental conjunction is in this regard similar to set-theoretic union. If sets A and B contain Todd, the union of A and B does not contain Todd twice-over.

Third, I’ll suppose that (IP) holds no matter whose ideas they are, human or divine. (IP) allows for infinitely complex mental conjunction. This will prove especially important when it comes to Section 10. In the human case, nothing about (IP) rules out indeterminate ideas or ideas with disjunctive conjuncts. (IP) does not imply that there is exactly one correct specification of which ideas are parts of the mental conjunction with the label of (say) ‘knowing that \( p \)’. Nor does it preclude ideas from having additional bells and whistles which account for why, for example, human subjects might judge that apples are more typical fruits than plums. (IP) doesn’t imply that we would be able to correctly specify all the conjuncts of an idea either. Quite simply, (IP) does not usher in controversial theses about conceptual analysis.
5.3 Parts of Ideas

In this section, I offer a formal theory of ideational parthood. In the last section, I claimed that an idea is part of an idea just in case the first is identical to the second or to one or some of the second’s conjuncts. I’ll let this notion of ‘part’ serve as the chosen mereological primitive used to define the other mereological notions. Since the chosen primitive notion is ultimately understood in terms of mental conjunction, every derivative mereological notion will also be analyzable into a claim about mental conjunction.

I’ll use the upper-case ‘I’ with lower-case subscripts (e.g., \( I_i \), \( I_j \), \( x \), \( y \), and \( n \), where the latter three variables range over the natural numbers) to range over ideas. I’ll begin with the definitions of proper parthood and overlap:

**Definition 1.** \( I_x \) is a proper part of \( I_y \) iff \( I_x \) is part of \( I_y \) and not identical to \( I_y \).

**Definition 2.** \( I_x \) and \( I_y \) overlap iff some \( I_z \) is part of \( I_x \) and also \( I_y \).

Using Definitions 1 and 2, we define two other mereological notions:

**Definition 3.** \( I_x \) and \( I_y \) are disjoint iff they do not overlap.

**Definition 4.** \( I_y \) is a sum of the \( I_x \)s iff ((each of the \( I_x \)s is part of \( I_y \)) and (any \( I_z \) overlaps one of the \( I_x \)s iff \( I_z \) overlaps \( I_y \))).

Definitions 1–4 ultimately hang on the notion of parthood in (IP), which is itself understood in terms of mental conjunction. So each of Definitions 1–4 might also be understood as claims about mental conjunction. Definition 1 and (IP) together imply that an idea is a proper part of another when the first is one of the second’s conjuncts. Definition 2 and (IP) imply that ideas overlap when they’re the same idea or when an idea is a conjunct of both. Definition 3, along with Definition 2 and (IP), implies that two ideas are disjoint when they have no conjuncts in common.

Definition 4 deserves a little more reflection than the others. Consider the right-hand side of the biconditional: (each of the \( I_x \)s is part of \( I_y \)) and (any \( I_z \) overlaps one of the \( I_x \)s iff \( I_z \) overlaps \( I_y \)).

The left conjunct ensures that an idea is a sum of its parts, not of any non-parts. The right conjunct is
another biconditional: any $I_z$ overlaps one of the $I_x$s iff $I_z$ overlaps $I_y$. The left-to-right portion of this biconditional (i.e., if any $I_z$ overlaps one of the $I_x$s then $I_z$ overlaps $I_y$), rules out any scenario in which an idea overlaps one of the $I_x$s of which $I_y$ is a sum but doesn’t overlap $I_y$ itself. The right-to-left portion (i.e., if any $I_z$ overlaps $I_y$, then $I_z$ overlaps one of the $I_x$s) rules out scenarios in which an idea overlaps the sum but not any of the ideas of which it is a sum.

With the definitions in tow, we may turn to the axioms. The first says that ideational proper parthood is asymmetric:

**Asymmetry.** If $I_x$ is a proper part of $I_y$, $I_y$ isn’t a proper part of $I_x$.

Suppose, for instance, that the idea of being crimson is a conjunction of ideas whose conjuncts include, among other ideas, the idea of being red. Since being crimson is conjunction of ideas, whatever satisfies being crimson will satisfy each and every one of its conjuncts. So each of those conjuncts specifies only a portion of what it takes for something to be crimson. This is intuitively right: it takes more for something to be crimson than it does to be red because being crimson is a specific way of being red. These considerations, I believe, underlie the judgment that the idea of being red is part of but not identical to the idea of being crimson. For the same reasons, being crimson is not a proper part of being red. Since it takes more to be crimson than it does to be red, the idea of being red does not contain whatever more it takes to be crimson. The relation ‘is one among other mental conjuncts of’ is asymmetric. So proper parthood is, too.

The second axiom says that ideational proper parthood is transitive:

**Transitivity.** If $I_x$ is a proper part of $I_y$, and $I_y$ is a proper part of $I_z$, then $I_x$ is a proper part of $I_z$.

Let’s suppose that the idea of being red isn’t simple. One plausible candidate for one of its parts is the idea of being colored. Now if the idea of being colored is a mental conjunct of the idea of being red, and the idea of being red is a mental conjunct of the idea of being crimson, it would appear that the idea of being colored is a mental conjunct of the idea of being crimson. An idea, which is a mental conjunction, cannot be a mental conjunct of another idea unless it carries all of its conjuncts with it as conjuncts of the larger mental conjunction. Unlike the relation of set-membership, the ‘is a mental conjunct of’ relation is transitive. Consequently, proper parthood is transitive, too.
The third axiom says that a proper part of an idea is always accompanied by another, disjoint proper part:

**Weak Supplementation.** If $I_x$ is a proper part of $I_y$, then $I_y$ has another proper part disjoint from $I_x$.

No mental conjunction has only one conjunct. A conjunction of ideas requires at least two. Furthermore, no conjunct appears twice in a mental conjunction (see the second to last paragraph of the previous section). With this in mind, let’s conduct a short exercise. Of an arbitrary mental conjunction, pick one of its conjuncts (which may be a mereologically simple idea having no conjuncts of its own or not), call it “the selection” and call the remaining conjuncts “the remainder.” The selection includes whatever mental conjuncts, if any, our pick includes. Since no idea is a conjunct of both the selection and the remainder, neither the selection nor the remainder are parts of each other. Nor do they share a common conjunct. Therefore, if an idea $I_x$ is one of $I_y$’s conjuncts, $I_y$ has another conjunct, $I_z$, which has no conjuncts in common with $I_x$. Given (IP), then, if $I_x$ is a proper part of $I_y$, then $I_y$ has another proper part disjoint from $I_x$.

Logical conjunction allows for infinitely long conjunctions as well as inconsistent conjunctions. A mental conjunction may also be infinitely complex or inconsistent. If there is an infinite mind with infinite cognitive resources, we may suppose it forms every possible mental conjunction, including all inconsistent conjunctions, such as *being an x such that x is red and not red*. Given an infinite mind, it seems reasonable to suppose that there is a mental conjunction of any specifiable set of ideas:

**General Sum Principle.** Given a specifiable set of ideas, there is a sum of those ideas.

Assuming that there’s an infinite mind, given any specifiable set of propositional ideas or monadic or relational ideas, there is a sum of them. Putting it this way does not license the existence of a paradoxical sum which is an idea of all ideas. We may distinguish an idea which every idea satisfies from an idea which successfully “collects” all ideas as an idea of them. The idea of being an idea is an idea that every idea satisfies, even itself. This is not paradoxical. There is also a collective idea of this, that, and the other—an idea which gathers the things in an idea of them. Suppose we attempt to define a set which collects all ideas. Since every idea satisfies the idea of being an idea, our attempted definition may use the idea of being an idea as the entrance condition to that set. There
is no such specifiable set. For as soon as we have a purported case of an all-inclusive collecting idea, the idea of that set will not be a member of the defined set. The General Sum Principle does not license the existence of a universal collecting idea for there is no such specifiable set of all ideas.\textsuperscript{13}

If there is no infinite mind, then the prospects of the General Sum Principle seem dim. Without an infinite mind, there is little reason to think that every possible combination of ideas forms a sum, especially if we construe sums of ideas as ideas themselves. If there is no infinite mind, then which sums exist may be constrained by the linguistic practices of communities and the mental lives of individual thinkers.

The next two axioms govern ideas of ideas. There are ideas like the idea of being a cat or the idea of Fred. But there are also ideas of ideas, like the idea of being the idea of being a cat and the idea of being the idea of Fred. When an idea is an idea of another, I’ll call the idea of an idea a \textit{meta-idea} and the idea it is an idea of its \textit{base idea}. This distinction is relative, of course, since meta-ideas may serve as base ideas for further meta-ideas. Now suppose the idea of being a cat has, as a part, the idea of being mammalian. Then, it would seem that the meta-idea \textit{being the idea of being a cat} would have, as a part, the idea of having being mammalian as a part. Hence, some of a meta-idea’s parts concern which ideas are parts of its base idea. Let us add to the mereology, then, this further axiom:

\textbf{Inclusivity}: If \(I_x\) is part of \(I_y\), then the idea of having \(I_x\) as a part is part of the further idea of being \(I_y\).

Now if a meta-idea has parts concerning which ideas are parts of its base idea, we might also think they have parts which inform us about which ideas are not parts of their base ideas. The idea \textit{being a dog}, for example, isn’t part of \textit{being human}. To be the idea of being human precludes its having \textit{being a dog} as a part. So perhaps the meta-idea \textit{being the idea of being human} has, as a part, the idea \textit{being an x such that x doesn’t have the idea of being a dog as a part}, or, less precisely, the idea of not having \textit{being a dog} as part. If meta-ideas behave this way, the following principle holds:

\textbf{Exclusivity}: If \(I_x\) isn’t part of \(I_y\), then the idea of not having \(I_x\) as a part is itself part of the further idea of being \(I_y\).

\textsuperscript{13}This follows more immediately if sets are ideas, as I argue in Chapter 5. Menzel (1987) argues convincingly that if sets are ideas, no paradox results.
Principles relating to ideational parthood will serve as conditions on frames which validate various modal axioms. In the next two sections, I turn my focus to other core components of the semantics.

5.4 Worldhood

If there is an infinite mind, we may consistently suppose it has an idea of whatever exists. Given a mind with an infinite store of cognitive resources, neither the number of things nor their complexity is too great to be thought. God’s mind, we may suppose, not only has ideas of you and me and each and every thing that exists. It also has an idea of everything taken together—an idea of the whole world.

An idea of the whole world—the nature of this idea depends on what the world is. One may adopt the interrelatedness conception of worldhood, which says that a world consists of everything its parts are spatiotemporally related to.14 On this picture, a world is an island universe and there may be more than one. Or one may adopt the the totality conception, which says that a world consists of everything that exists, including however many island universes there are.15 The formalism of my semantics is compatible with either view. If the actual world is an island universe, the idea of the actual world is an idea of the island universe. If the actual world is the totality of everything that exists, then an idea of the actual world is an idea of the totality of everything that exists.

Although worldhood is central to contemporary approaches to modality, some have doubted whether the concept or idea of worldhood is coherent at all. For example, Bas van Fraassen (1995) argues that ‘world’ isn’t a count-noun and that there’s no reason to believe it refers to anything. Related worries concern the status of absolutely unrestricted quantification. Perhaps there’s no such thing as a collection of everything that exists—or even if there is, maybe we can’t even refer to it.16 I shall assume that these worries are answerable, in principle. Since the concept of worldhood is central to both my approach (because I grant that there’s an idea of the actual world) and to various possible worlds approaches to modality (because these claim not only that there is an actual world but many

15 Stalnaker (1976, 69–70).
16 See, for example, many of the essays in Rayo and Uzquiano (2007).
more possible ones), neither kind of approach uniquely carries the burden of discharging worries about its coherence.

Call the actual world \textit{alpha}. What makes alpha alpha is the particular way it is, in every detail. Since the way alpha is involves the ways that you and I and everything else are, the idea of alpha involves ideas about you and me and everything else. Alpha is also a world in general, the idea of alpha also includes that it is a world in general. The idea of being a world in general is an abstraction from the idea of being \textit{this} world: every part of the idea of being a world in general is part of the idea of being this world but presumably the converse does not hold. The idea of being a world in general presumably doesn’t include every idea of how alpha is no more than the idea of being a man in general includes every idea of how Barack Obama is. But just as the idea of being a man in general is an idea any possible man would satisfy if he were actual, the idea of being a world in general is an idea any possible world would have satisfied if it had been actual.

\textit{Being a world} is an idea any possible world would satisfy if it had been actual. But does the vast array of possible worlds determine the content of the idea of being a world in general or does the content of being a world somehow determine what is common to all possible worlds? There are at least two ways for the array of possible worlds to determine the content of \textit{being a world}. First, one might construe ideas as sets of possibilia whose members would have satisfied the idea if they had been actual. Construing ideas this way would allow one to identify \textit{being a world} with the set of possible worlds. I reject this treatment of ideas in Section 2. Second, one might think the array of possible worlds determines the content of \textit{being a world} in the following way. God begins with the stock of possible worlds and grasps the commonalities among them to develop an idea of what it takes to be a world in general. Then, whether an idea is part of \textit{being a world} is grounded in whether each possible world, if actual, would have satisfied that idea.

To grasp the commonalities among all possible worlds, it is crucial that one begins with the stock of all and only possible worlds. Leaving out some possible world(s) might lead one to think that some feature is common to all worlds when it is not. Counting some non-possible world(s) as possible worlds might lead one to miss a feature which is common among all possible worlds. Therefore, to grasp what is common among possible worlds as what is common among possible worlds, one must first separate all and only the possible worlds from everything else in a principled way. To do this, one would have to know already what worldhood consists in. That is, one already
would have to grasp the idea of being a world in general before grasping the commonalities among all the possible worlds. Hence, what is common to all possible worlds does not ground the content of being a world.

Instead, the content of being a world grounds the commonalities among possible worlds. This is true whether we identify possible worlds with ideas of worlds or whether possible worlds are non-existent objects that those ideas represent. Suppose the first. The idea of any particular world is thereby a “possible world.” Given that the idea of any particular world is a mental conjunction, one of its conjuncts is an idea of being a world in general. Therefore, each idea of a particular world has the idea of being a world in general as a part. If parthood is transitive, then the parts of being a world are also parts of every idea of a world. If ideas of worlds are possible worlds, those commonalities among possible worlds result from their having being a world as a part. Possible worlds have a common core, the idea of being a world in general, and they differ in how they fill out from there.

Now suppose the second, that possible worlds are instead the non-existent intentional objects of the ideas which we previously identified with possible worlds. Ideas of worlds which are merely possible are ideas of worlds which do not exist. It is plausible, I think, that insofar as possible worlds in this sense have any being at all, they have it in virtue of being an object of an idea with that idea’s particular intentional or representational content. Then, if being an object of an idea of a possible world is responsible for the being of a possible world, then whatever features a possible world has would also seem to be determined by the content of the idea it is an object of. The idea of any particular world has being a world as a part. So each possible world, if it had been actual, would have satisfied being a world and hence all its parts. Even if possible worlds are not ideas but non-existent objects of ideas, I believe it is more plausible than not that the commonalities we find among possible worlds are determined by the content of the idea of being a world in general. For if any possible world had its being as a result of being the object of an idea of a possible world, and that idea’s content includes the idea of being a world in general, then the commonalities we find among possible worlds are determined by the content of the idea of being a world in general.

Either way, whether the idea of being an F is part of being a world is responsible for whether every possible world would have been F if it had been actual. God finds the commonalities he does among possible worlds precisely because each possible world, if actual, would have been a world and hence satisfied not only being a world but also all the ideas which compose it. If the array of
possible worlds fails to ground which ideas are parts of being a world, what does ground whether or not an idea is part of being a world? Ideas are not reducible to other entities like extensions or functions. Whether or not an idea exists plausibly depends on whether there is a thinker whose idea it is and also perhaps on, in the case of ideas of individuals, whether the object of the idea exists. But given the existence of an idea, its parts come along for free as its parts.

5.5 Propositional Ideas

God’s understanding of the actual world in its entirety consists in his grasp of the complete idea of the actual world. Speaking in terms of understanding is sometimes more natural and illuminative than speaking in terms of ideas, especially at the scale of the entire actual world. From now on I will frequently discuss the structure of God’s understanding as a roundabout way of discussing the structure of divine ideas. If Fred is actually blue, for example, then God’s understanding of the actual world includes an understanding of Fred’s being blue. And, generally, if a is actually F, then God’s understanding of the actual world includes an understanding of a’s being F. The locution ‘a’s being F’ is a participial nominalization of the sentence ‘a is F’. A sentence’s participial nominalization signifies an idea which represents the very state of affairs the sentence concerns. I’ll call ideas of this form propositional ideas.

The propositional idea of Fred’s being blue also has a mereological structure. To bring that structure into relief, let’s introduce some shorthand. Let [Fred] be Fred’s complete idea, and let [being blue] be the idea of being blue. So [Fred] captures God’s full understanding of the way Fred actually is, and [being blue] captures God’s full understanding of the property of being blue. Both ideas appear in the propositional idea of Fred’s being blue. God’s full understanding of Fred includes [being blue], given that Fred is actually blue. God’s understanding of Fred’s being blue consists in [Fred] having [being blue] as one of its parts.

Given the current conception of ideas as having parts, something x satisfies the whole idea if and only if x satisfies the ideas which compose the whole idea. [Fred] is a mental conjunction whose conjuncts include all the ideas that Fred actually satisfies. But [Fred] also represents what something would be like if it satisfied all of the ideas that compose [Fred]. If there is such a thing that satisfies all of the ideas that compose [Fred], then [Fred] represents that thing. Let’s suppose there is such a
thing—Fred. [Fred] represents Fred in all his glory—as tall, if he’s tall, as blue, if he’s blue, and so on. The propositional idea of Fred’s being blue, the idea which consists in [Fred]’s having [being blue] as a part, specifically focuses on one among many of Fred’s features—his blueness. The idea which consists in [Fred]’s having [being blue] as a part represents Fred specifically as being blue.

The propositional idea of Fred’s being blue involves [Fred], an idea which accounts for everything truly predicable of Fred. An individual’s complete idea is an idea whose parts are ideas which capture everything about that individual. If Fred is six feet tall, for example, then the idea of being six feet tall is part of Fred’s complete idea. If Fred isn’t six feet tall, being six feet tall isn’t part of Fred’s complete idea. Generally, a is F if and only if the idea of being F is part of a’s complete idea. Complete ideas are highly complex, perhaps infinitely so. And if a complete idea also contains how its individual is related to everything else, it will be structurally similar to a Leibnizian monad, which in itself mirrors the entire universe. Although we are surely unable to grasp ideas with such complexity, this is no issue for an infinite mind.

But there many other ideas of Fred, ideas of Fred which contain fewer ideas than the complete idea of Fred. If there are many ideas of Fred, then, presumably, there are many ideas of Fred available to play a role in a propositional idea about Fred. One candidate for such an idea is what I’ll call Fred’s individual idea. An individual idea’s parts not only capture but also determine what is essential to being that individual. If being a person is part of Fred’s individual idea, Fred is essentially a person. If being married isn’t part of his individual idea, then he’s not essentially married. And if being unmarried is not part of his individual idea then he’s not essentially unmarried either. If there is more than a nominal difference between an individual’s actual and essential features, then an individual’s complete idea will contain more than its individual idea.

Whether there is more than a nominal difference between an individual’s actual and essential features is the subject of a dispute between Leibniz and Arnauld. In their correspondence, they agree that there is a complete concept of every individual, a concept that contains everything that ever happens to that individual. They also agree that there is an individual concept for each individual, a concept which contains concepts of the essential features of the individual. They disagree about whether individual concepts are complete concepts, i.e., whether the concept of what it takes to be a particular individual contains everything that ever happens to that individual. Leibniz believes “the individual concept of each person contains once and for all everything that will ever happen to
For Leibniz, then, any actual property is an essential property. Caesar crosses the Rubicon, for instance, so Caesar’s individual concept contains the concept of crossing the Rubicon.\(^{18}\) Arnauld disagrees:

> Since it is impossible that I should not always have remained myself, whether I had married or lived in celibacy, the individual concept of myself contained neither of these two states; just as it is well to infer: this block of marble is the same whether it be at rest or be moved; therefore neither rest nor motion is contained in its individual concept (LA 30)

Arnauld plausibly suggests that individual concepts are incomplete. Neither marrying nor not marrying is part of his own individual concept. Neither is part of what it takes to be Arnauld: he might have married or might not have. Arnauld distinguishes the complete concept, which contains everything about that individual, from the individual concept, which contains less than the complete concept and only contains what is essential to being Arnauld.

Like both Arnauld and Leibniz, I will insist that there is at least a nominal distinction between individual and complete ideas. If individual and complete ideas were identical, we’d be left with the Leibnizian view that everything that ever happens to an individual is essential to it. The formalism underlying my approach to quantified modal logic is compatible with this Leibnizian view. But it is also compatible with the view that individual ideas are proper parts of and therefore different from complete ideas.

The nominal distinction between individual and complete ideas is important within the semantics because whether the meaning of a name is an individual or complete idea depends on whether the name appears within the scope of a modal operator. If the context is whether it is necessary or possible that Fred is tall, Fred’s individual idea is in play. If the context is whether it is true that Fred is tall, his complete idea is in play. Hence, there is not just one propositional idea of Fred’s being tall. There is the propositional idea which consists in Fred’s individual idea having the idea of being tall as a part and the propositional idea which consists in Fred’s complete idea having the idea of being tall as a part.

\(^{17}\)LA 12.

5.6 Truth

We now have the raw materials to construct an idealist semantics for quantified modal logic without an accessibility relation among possible worlds or world-like stand-ins. In the next several sections, I present the core concepts of the semantics.

It’s true that an actual individual \( a \) is \( F \) just in case the actual world is such that \( a \) is \( F \). And the actual world is such that \( a \) is \( F \) just in case the propositional idea of \( a \)’s being \( F \) is part of the idea of the actual world. So it’s true that \( a \) is \( F \) just in case the idea of \( a \)’s being \( F \) is part of \( A \), the idea of the actual world:

\[
\text{\begin{figure}[h]
\centering
\includegraphics[width=0.3\textwidth]{figure.png}
\caption{The idea of \( a \)’s being \( F \) is part of the idea of the actual world.}
\end{figure}}
\]

\( A \) is a mental conjunction whose parts include propositional ideas. Since these propositional ideas are parts of \( A \), they represent how the actual world is. Hence, where \( \phi \) is a wff, \( [\phi] \) is the propositional idea of being such that \( \phi \), \( A \) is the idea of alpha, and ‘;’ reads “is part of”:

\[
(A) \ \phi \text{ is true } =_{df} [\phi] \; ; A.
\]

Since \( A \) represents how the actual world is in all its complexity, someone who grasps this idea has a full understanding of the actual world. Such an understanding includes ideas of every actual individual and also the ways they’re related to one another.

Since I aim to offer an applied semantics for S5, which contains classical logic, I will assume the actual world behaves classically in ways that secure the theorems of classical logic.\(^{19}\) I assume, first,

\(^{19}\)I’ve refashioned much of what remains in this section from Chapter 2.
that $\mathcal{A}$, the idea of the actual world, is *complete*: for every pair of propositions $\phi$ and $\neg\phi$, at least one of either $[\phi]$ or $[\neg\phi]$ is part of $\mathcal{A}$. I also assume that $\mathcal{A}$ is *consistent*: for every pair of propositions $\phi$ and $\neg\phi$, at most one of either $[\phi]$ or $[\neg\phi]$ is part of $\mathcal{A}$. Therefore, if $[\phi]$ is a part of $\mathcal{A}$, $[\neg\phi]$ isn’t, and *vice versa*. $\mathcal{A}$’s completeness secures the law of excluded middle and guarantees that there are no truth-value gaps. Its consistency secures the law of non-contradiction and guarantees that there are no true contradictions.

Now (A) implies that an atomic formula such as ‘$F_a$’ is not true when $[F_a]$ is not part of $\mathcal{A}$. But it also implies that ‘$\neg F_a$’ is true when $[\neg F_a]$ is part of $\mathcal{A}$. Where $a$ is an actual individual, $[F_a]$ is not part of $\mathcal{A}$ just in case $[\neg F_a]$ is part of $\mathcal{A}$. Let’s consider each direction of this biconditional. First, left-to-right. If a full understanding of the actual world does not include $a$’s being $F$, then the actual world is such that $a$ is not $F$. If the actual world is such that $a$ is not $F$, then there’s an actual individual $a$ which is not $F$. And if an actual individual $a$ is not $F$, then a full understanding of the actual world includes $a$’s not being $F$. Second, consider right-to-left. If a full understanding of the actual world includes that $a$’s not being $F$, then an actual individual $a$ is not $F$. If an actual individual $a$ is not $F$, then the actual world is such that $a$ is not $F$. And if the actual world is such that $a$ is not $F$, then a full understanding of the actual world does not include $a$’s being $F$. Hence, as long as $\mathcal{A}$ is both complete and consistent, $[\neg F_a]$ is part of $\mathcal{A}$ if and only if $[F_a]$ is not part of $\mathcal{A}$.

Other auxiliary assumptions about parthood relations among ideas in $\mathcal{A}$ seem to follow from our intuitive understanding of ‘...and...’ statements, ‘...or...’ statements, ‘if... then...’ statements and the like:

$$(\mathcal{A}\supset) \text{ if } [\phi \supset \psi] \in \mathcal{A} \text{ and } [\phi] \in \mathcal{A}, \text{ then } [\psi] \in \mathcal{A}.$$ 

$$(\mathcal{A}\neg) \text{ if } [\phi] \in \mathcal{A} \text{ iff } [\neg\phi] \in \mathcal{A},$$

$$(\mathcal{A}\land) \text{ if } [\phi \land \psi] \in \mathcal{A} \text{ iff } [\phi] \in \mathcal{A} \text{ and } [\psi] \in \mathcal{A},$$

$$(\mathcal{A}\lor) \text{ if } [\phi \lor \psi] \in \mathcal{A} \text{ iff either } [\phi] \in \mathcal{A} \text{ or } [\psi] \in \mathcal{A}. $$

These assumptions secure theorems and validate various inferences of first-order logic. For example, if $[\phi \land \psi]$ is part of $\mathcal{A}$, then the conjunction $\phi \land \psi$ is true, by (A). But $[\phi \land \psi]$ is part of $\mathcal{A}$ if and

\[\text{20} \text{ Also, if either } [\neg F_a] \text{ is part of } \mathcal{A} \text{ or } [F_a] \text{ is not part of } \mathcal{A}, \text{ then } [\text{being an } F] \text{ is not part of } [a].\]
only if both $[\phi]$ and $[\psi]$ are, by $(A \land)$. So if $[\phi \land \psi]$ is part of $A$, then so are $[\phi]$ and $[\psi]$. If $[\phi]$ and $[\psi]$ are both parts of $A$, then both $\phi$ and $\psi$ are true, by (A). Hence, $(A \land)$ validates the inferences from a conjunction to each conjunct.

5.7 **Necessity**

Let $W$ be the idea of being a world in general. When $W$ includes the idea of an individual’s being such-and-such, then that individual’s being such-and-such is a precondition of worldhood. If that individual’s being such-and-such is a precondition of worldhood, it is necessary that the individual is such-and-such. More generally, a proposition is necessary when its corresponding propositional idea is part of $W$:

$$(N) \quad \square \phi \text{ is true } =_{df} \ [\phi] \ i W.$$

The diagram below depicts the necessity of the proposition that $a$ is F.

![Figure 5.2: The idea of $a$’s being F is part of the idea of being a world.](image)

We may also define possibility in terms of necessity. Something is possible when it isn’t necessary that it isn’t the case, i.e., $\Diamond \phi$ is true if and only if $\neg \square \neg \phi$ is true. $\neg \square \neg \phi$ is true when $\neg \phi$ isn’t part of $W$. Therefore:

$$(P) \quad \Diamond \phi \text{ is true } =_{df} \ [\neg \phi] \not\in W.$$
This makes intuitive sense. If \( \neg \phi \) were part of \( \mathcal{W} \), it would be necessary that \( \neg \phi \). That is, \( \phi \) would be impossible—i.e., not possible. What is not not possible is possible, so by negating the feature in the semantics that would make \( \phi \) not possible (\( [\neg \phi] \)'s being part of \( \mathcal{W} \)) we affirm \( \phi \)'s possibility.

Like the propositional parts of \( \mathcal{A} \), I assume (reasonably, I think) that the propositional parts of \( \mathcal{W} \) also behave in ways that secure our intuitive understanding of ‘...and...’ statements, ‘...or...’ statements, ‘if... then...’ statements and so on. This suggests a number of principles that govern how propositional properties operate in \( \mathcal{W} \), including:

\((\mathcal{W} \supset)\) if \([\phi \supset \psi]\) and \([\phi]\) are parts of \( \mathcal{W} \), so is \([\psi]\),

\((\mathcal{W} \neg)\) \([\phi]\) is part of \( \mathcal{W} \) iff \([\neg \neg \phi]\) is,

\((\mathcal{W} \wedge)\) \([\phi \wedge \psi]\) is part of \( \mathcal{W} \) iff \([\phi]\) and \([\psi]\) are, and

\((\mathcal{W} \lor)\) if \([\phi]\) is part of \( \mathcal{W} \), so is \([\phi \lor \psi]\).\(^{21}\)

Joined with (N), these principles validate a number of intuitive modal inferences. For instance, it can be easily shown that \((\mathcal{W} \wedge)\) and (N) validate the inferences from \( \Box(\phi \wedge \psi) \) to both \( \Box \phi \) and \( \Box \psi \) and from both of these back again to \( \Box(\phi \wedge \psi) \). I’ll demonstrate the left-to-right portion of this biconditional. If \([\phi \wedge \psi]\) is part of \( \mathcal{W} \), then \( \Box(\phi \wedge \psi) \) is true, by (N). If \([\phi \wedge \psi]\) is part of \( \mathcal{W} \), so are \([\phi]\) and \([\psi]\), by \((\mathcal{W} \wedge)\). If \([\phi]\) and \([\psi]\) are parts of \( \mathcal{W} \), then \( \Box \phi \) and \( \Box \psi \) are true, by (N). Therefore, if \( \Box(\phi \wedge \psi) \) is true, so are \( \Box \phi \) and \( \Box \psi \).

We now have a crude picture of truth, necessity, and possibility. In the next section, I present the semantics more formally.

### 5.8 Names

The semantic value of a name is not its referent but instead an idea of the referent. I’ll sometimes say that some name “denotes” an idea, but I only mean that the interpretation function assigns the idea as the semantic value of the name, not that the idea is the referent of the name. For now it will be beneficial to think of this denoting relation as a two-part function which takes us from a name to its referent and then from its referent to an idea of the referent.

\(^{21}\)The biconditional that \([\phi \lor \psi]\) is part of \( \mathcal{W} \) iff either \([\phi]\) or \([\psi]\) is, presumably doesn’t hold for \( \mathcal{W} \) because \( \mathcal{W} \) isn’t complete like \( \mathcal{A} \) is.
The official treatment is a bit more complicated than this, however. Let an idea of an individual be any idea which contains the idea of being identical to that individual. \( A \) is an idea of what is actual, so if \( A \) contains a propositional idea of something’s being (or not being) such-and-such, then \( A \) implicates the idea of that thing as an idea of an actual individual. Of all the ideas of an individual \( \alpha \) implicated in \( A \), let the *largest* of them be the one which contains each of them as a part. Since \( A \) is both complete and consistent, the largest ideas of actual individuals happen to be the complete ideas of actual individuals.

Again, an idea of an individual is any idea which contains the idea of being identical to that individual. If there’s an idea of an individual which is itself a constituent of a propositional idea which isn’t part of \( W \), then \( W \) implicates the idea of that thing as an idea of an individual which is possible. \( W \) is an idea which determines what’s necessary, so if \( W \) contains a propositional idea of something’s being (or not being) such-and-such, then \( W \) requires that the individual be such-and-such (or not be such-and-such). Of all the ideas of an individual implicated in \( W \) as being this way or that, let the *largest* of them be the one which contains each of them as a part. Let’s call these largest ideas of individuals implicated in \( W \) *individual ideas*.

There are both individual and complete ideas of individuals, and whether a name denotes one or the other depends on whether the name occurs within the scope of a modal operator. For instance, in a non-modal context concerned with the truth of whether Fred is blue, ‘Fred’ denotes Fred’s complete idea, \([\text{Fred}_c]\). In a modal context concerned with the truth of whether Fred is possibly or necessarily blue, ‘Fred’ denotes Fred’s individual idea, \([\text{Fred}_i]\). So ‘Fred is blue’ is true just in case the propositional idea which consists in \([\text{Fred}_i]\)’s having [being blue] as a part is itself part of \( A \). And ‘Fred is blue’ is necessarily true just in case the propositional idea which consists in \([\text{Fred}_i]\)’s having [being blue] as a part is itself part of \( W \).

\([\text{Fred}_c]\) is a mental conjunction of all the ideas which reflect how Fred truly is. \([\text{Fred}_i]\) is a mental conjunction of all the ideas which reflect how Fred must be, if he is at all. Given that whatever way Fred must be is some way he truly is, every part of \([\text{Fred}_i]\) is part of \([\text{Fred}_c]\). If [being blue] is part of \([\text{Fred}_i]\), it is part of \([\text{Fred}_c]\). Therefore, if the propositional idea which consists in \([\text{Fred}_i]\)’s having [being blue] as a part is itself part of \( A \), then the propositional idea which consists in \([\text{Fred}_c]\)’s having [being blue] as a part is also part of \( A \). This holds generally, too: if the propositional idea which consists in \([a_i]\)’s having [being blue] as a part is itself part of \( A \), then the propositional idea which consists in \([a_c]\)’s having [being blue] as a part is also part of \( A \).
which consists in \([a, c]\)’s having [being blue] as a part is also part of \(A\). I’ll call this the Individual-Completeness Principle, or the I-C Principle for short. The I-C principle will be important later on when we turn to the more formal details of the semantics.

5.9 Predicates

A predicate’s semantic value is not its extension in the actual world.\(^{22}\) Nor is a predicate’s semantic value its extension across possible worlds, whether construed as a set of possible instances or as a function defined over possible worlds. Here, a predicate’s semantic value is an idea of what the predicate denotes or purports to denote, namely some property. In standard treatments of predicates, each \(n\)-ary predicate receives a single semantic value in virtue of which one may calculate the truth-value of any atomic formula in which the predicate appears. This is not what happens here. I’ll explain what happens instead in three steps.

First, for any \(n\)-ary atomic sentence, there is a set of \(n\) abstractions, one for each argument place. Given the atomic sentence \(F_a\), there is the single lambda-abstraction \(\lambda x F x\). Given \(G_{ab}\), there are two different lambda-abstractions, \(\lambda x G x b\) and \(\lambda x G a x\). Given \(H_{abc}\), there are three different lambda-abstractions: \(\lambda x H x b c\) and \(\lambda x H a x c\), and \(\lambda x H a b x\). And so on. There is a function from atomic sentences to sets of their corresponding lambda-abstractions.

Second, lambda-abstractions are often taken to denote properties which are treated set-theoretically as functions from (possible) individuals to truth values. In Montague semantics, for example, ‘\(\lambda x F x\)’ denotes the property of being an \(x\) such that \(x\) is \(F\), which is then treated as a function from possible worlds to a further function from possible individuals to truth-values. Here, lambda-abstractions denote ideas, not properties. ‘\(\lambda x F x\)’, for example, denotes the idea of being an \(x\) such that \(x\) is \(F\) or, more simply, the idea of being an \(F\). I treat these ideas as primitives in the semantics, not as sets or functions.

Third, lambda-abstractions are notoriously difficult to read, so I will replace the standard formalism of the lambda-calculus with a simpler formalism designed for my simpler purposes. An abstraction from the \(m^{th}\) argument place in a formula encloses the formula in brackets and replaces

\(^{22}\)Much of this material is translated from talk about properties as semantic values in Chapter 3. I left the nature of those properties open wide enough to later identify them with ideas. The properties that now feature as the objects of ideas are different in kind from the properties that I now identity as ideas.
the $m^{th}$ term with `$x$'. Given ‘Fa’, there is the abstraction `[Fx]’, which denotes the idea of being an F. Given ‘Gab’, we have `[Gx]’, which denotes the idea of being an $x$ such that $a$ is G-related to $x$, and `[Gxb]’, which denotes the idea of being an $x$ such that $x$ is G-related to $b$.

Lambda-abstraction is also useful for representing propositional ideas. Ordinarily, $\lambda x \phi$ denotes the propositional property of being such that $\phi$. Not here. Instead, it denotes the propositional idea of $\phi$, $[\phi]$. I’ll represent the atomic propositional idea of $a$’s being $F$ as $[[Fx] ; [a_c]]$. Thus, ‘[Fa]’ and ‘[[Fx] ; [a_c]]’ denote the same propositional idea of $a$’s being $F$.

5.10 Negation

How does the current semantics account for negation? According to the account of truth in Section 6, ‘Fa’ is true when the propositional idea $[Fa]$ is part of $\mathcal{A}$. And according to the account of propositional ideas in Section 5, the propositional idea $[Fa]$ is the idea of a’s being $F$, the idea which consists of the idea of a having the idea of being $F$ as a part. Given that ‘Fa’ doesn’t occur within a modal operator, Sections 8 says that the idea of a involved in $[Fa]$ is $[a_c]$, a’s complete idea. And Section 9 says that $[Fx]$ is the idea of being $F$. Put all these pieces together, and we get the following as an account of the truth of a non-modal formula: where ‘¡’ signifies the ‘is part of’ relation, ‘Fa’ is true just in case $[[Fx] ; [a_c]] ; \mathcal{A}$. That is, ‘Fa’ is true when the propositional idea which consists in $[a_c]$ having $[Fx]$ as a part is itself part of $\mathcal{A}$. Since negation is usually defined in terms of a sentence’s being not true, this treatment of truth and the current account of propositional ideas suggests a few possible treatments of negation.

First, we might say that ‘$\neg$Fa’ is true when $[[Fx] ; [a_c]]$ is not part of $\mathcal{A}$. If a full understanding of alpha fails to include an understanding of a’s being $F$, then alpha is certainly not such that $a$ is $F$. And if alpha isn’t such that $a$ is $F$, then $a$ is not $F$. Conversely, if $a$ is not $F$, then a full understanding of alpha does not include an understanding of a’s being $F$. And if a full understanding of alpha doesn’t include an understanding of a’s being $F$, then $[[Fx] ; [a_c]]$ is not part of $\mathcal{A}$. This strategy gets the right results in such a simple case, but I will not take this strategy.

Second, we might say that ‘$\neg$Fa’ is true when $[\neg Fa]$ is part of $\mathcal{A}$. Then, the propositional idea $[\neg Fa]$ would be understood as the idea which consists in $[a_c]$’s not having $[Fx]$ as a part. This second strategy also gets the right result in this simple case. If a full understanding of alpha includes an
understanding of a’s not being F, then alpha is such that a is not F. And if alpha is such that a is not F, then a is not F. Conversely, if alpha is such that a is not F, then a full understanding of alpha includes an understanding of a’s not being F. And if a full understanding of alpha includes an understanding of a’s not being F, then the propositional idea \([F x \not\in [a_c]]\) is part of \(A\). I will not take this strategy either.

Third, we might say as before that ‘\(\neg F a\)’ is true when \([\neg F a]\) is part of \(A\), only this time with a different analysis of \([\neg F a]\). The difference between the previous and current strategies lies in whether \([a_c]\) lacks an idea as a part or instead has a negated idea as a part. Before, we understood \([\neg F a]\) as the propositional idea which consists in \([a_c]\)’s merely not having \([F x]\) as a part. Instead, we might understand \([\neg F a]\) as the propositional idea which consists in \([a_c]\)’s having \([\neg F x]\) as a part. Again, this strategy gets the right results in such a simple case. If a full understanding of alpha includes an understanding of a’s being not F, then alpha is certainly such that a is not F. And if alpha is such that a is not F, then a is not F. Conversely, if a is not F, then a full understanding of alpha includes an understanding of a’s being not F. And if a full understanding of alpha includes an understanding of a’s being not F, then \([\neg F x \mid [a_c]]\) is part of \(A\). This is not my preferred strategy either.

Finally, we may analyze the truth of ‘\(\neg F a\)’ in a way that falls somewhere between the second and third strategies. Like the second and third strategy, the chosen strategy implies that ‘\(\neg F a\)’ is true when \([\neg F a]\) is part of \(A\). Unlike the second strategy, the analysis of the true propositional idea \([\neg F a]\) does not merely consist in \([a_c]\)’s not having \([F x]\) as a part, though it does imply this. But unlike the third strategy, the analysis of the true propositional idea \([\neg F a]\) does not consist in \([a_c]\)’s having \([\neg F x]\) as a part either, though the chosen analysis is implied by that. Given that \([a_c]\) is a complete and consistent idea of how a is, a’s having \([\neg F x]\) as a part would preclude its having \([F x]\) as a part. For our assumptions in Section 6 imply that no complete idea of an actual individual includes two inconsistent ideas. An idea \(x\) precludes an idea \(y\) when the mental conjuction of \(x\) and \(y\) represents some impossible individual(s) or state(s) of affairs. For example, \(x\) may be the idea of being a blue tiger and \(y\) may be the idea of being red all over. The mental conjuction of the two represents an impossible individual since no individual could be blue and red all over. Of course, if \([\neg F x]\) were part of \([a_c]\), then \([a_c]\) would preclude \([F x]\). But as the blue tiger example suggests, we needn’t assume that when \([a_c]\) precludes \([F x]\), \([a_c]\) therefore has \([\neg F x]\) as a part. Even so, if \([a_c]\) precludes \([F x]\), then \([F x]\) isn’t part of \([a_c]\).
An idea’s precluding \([Fx]\) falls in between its having \([\neg Fx]\) as a part and its not having \([Fx]\) as a part. Why should we prefer this fence-sitting treatment of negation? First, it leaves open the possibility that \([a_i]\)’s having \([\neg Fx]\) as a part accurately characterizes the propositional idea \([\neg Fa]\). Second, and more importantly, treating negation in terms of preclusion predicts the sort of behavior we would expect from an individual’s individual and complete idea. If for whatever reason an individual idea \([a_i]\) precludes an idea such as [being \(F\)], any idea which has \([a_i]\) as a part will also preclude [being \(F\)]. Therefore, since the individual idea \([a_i]\) is part of the complete idea \([a_c]\), \([a_c]\) precludes any ideas that \([a_i]\) precludes. Tom’s individual idea precludes the idea of being blue when the mental conjunction of Tom’s individual idea and being blue represent something impossible. That may be because Tom’s individual idea has the idea of being not blue as a part or because there is only one color for Tom to be and Tom’s individual idea already represents him as being red, let’s say. If the mental conjunction of Tom’s individual idea and the idea of being blue already represent something impossible, then adding more ideas to the conjunction will not make that impossibility disappear. If Tom’s individual idea precludes the idea of being blue, so does his complete idea.

If there’s a propositional idea part of \(\mathcal{A}\) which consists in \([a_i]\)’s precluding some idea, then there’s a propositional idea which is also part of \(\mathcal{A}\) and consists in \([a_c]\)’s precluding that same idea. In Section 8, I introduced the Individual-Completeness Principle: if the propositional idea which consists in \([a_i]\)’s having [being \(F\)] as a part is itself part of \(\mathcal{A}\), then the propositional idea which consists in \([a_c]\)’s having [being \(F\)] as a part is also part of \(\mathcal{A}\). Treating negation as preclusion provides a similar principle. I’ll call the conjunction of both principles the Individual-Completeness Principles. Where \(\not\prec\) is the symbol for the preclusion relation, we may represent the principles in the following way:

1. If \([Fx] \not\prec [a_i]\) is involved in any propositional idea which is itself part of \(\mathcal{A}\), then the propositional idea which results from replacing \([Fx] \not\prec [a_c]\) for \([Fx] \not\prec [a_i]\) is itself part of \(\mathcal{A}\).

2. If \([Fx] \not\prec [a_i]\) is involved in any propositional idea which is itself part of \(\mathcal{A}\), then the propositional idea which results from replacing \([Fx] \not\prec [a_c]\) for \([Fx] \not\prec [a_i]\) is itself part of \(\mathcal{A}\).

I’ll call the conjunction of (1) and (2) the \(I\text{-}C\) Principles for short. The I-C principles will reappear later.
5.11 Quantification

The domain of quantification houses the meanings which the interpretation assigns to terms. Here the domain houses each individual’s complete and individual idea. Let the domain \( D \) contain ordered pairs of the form \( \langle [u_c], [u_i] \rangle \), where \([u_c]\) is a complete idea and \([u_i]\) is an individual idea of the same individual. Since the meaning of a name in non-modal contexts is a complete idea, a proposition about everything in such contexts concerns all the complete ideas \([u_c]\) such that \( \langle [u_c], [u_i] \rangle \in D \). If \([\text{Tom}_c], [\text{Tim}_c], \text{and } [\text{Terry}_c] \) exhaust the complete ideas in \( D \), then the proposition that everything is F is true just in case \( [[F x] ; [\text{Tom}_c]], [[F x] ; [\text{Tim}_c]], \text{and } [[F x] ; [\text{Terry}_c]] \) are all parts of \( A \).

The meaning of a name in modal contexts is an individual idea. So when a universal quantifier appears within the scope of a modal operator (“necessarily, everything...”), it concerns all the individual ideas \([u_i]\) such that \( \langle [u_c], [u_i] \rangle \in D \). If \([\text{Tom}_i], [\text{Tim}_i], \text{and } [\text{Terry}_i] \) exhaust the individual ideas in \( D \), then the proposition that necessarily, everything is F is true just in case \( [[F x] ; [\text{Tom}_i]], [[F x] ; [\text{Tim}_i]], \text{and } [[F x] ; [\text{Terry}_i]] \) are all parts of \( W \). Hence, the current semantics treats a quantifier differently depending on whether it occurs within the scope of a modal operator.

This treatment of the quantifiers requires temporary denotations to variables.\(^\text{\footnotesize 23}\) A non-modal formula like ‘\( \forall x F x \)’ is true iff for any \([u_c]\) such that \( \langle [u_c], [u_i] \rangle \in D \), ‘\( F x \)’ is true when \( x \) temporarily denotes the idea \([u_c]\). For when ‘\( x \)’ temporarily denotes \([\text{Tom}_c]\), then \( F x \) is true when \( [[F x] ; [\text{Tom}_c]] \) is part of \( A \). But if ‘\( F x \)’ is true no matter which \([u_c]\) such that \( \langle [u_c], [u_i] \rangle \in D \) ‘\( x \)’ temporarily denotes, then ‘\( \forall x F x \)’ is true. Similarly, a modal formula like ‘\( \Box \forall x F x \)’ is true iff for any \([u_i]\) such that \( \langle [u_c], [u_i] \rangle \in D \) ‘\( F x \)’ is true when \( x \) temporarily denotes the idea \([u_i]\). For when ‘\( x \)’ temporarily denotes \([\text{Tom}_i]\), then \( F x \) is true when \( [[F x] ; [\text{Tom}_i]] \) is part of \( W \). But if ‘\( F x \)’ is true no matter which \([u_i]\) ‘\( x \)’ temporarily denotes, then ‘\( \Box \forall x F x \)’ is true. Hence, a temporary assignment of ideas as the denotations of variables assigns an ordered pair \( \langle [u_c], [u_i] \rangle \) to each variable, and then whether the variable denotes the complete idea \([u_c]\) or the individual idea \([u_i]\) depends on whether or not the variable occurs within the scope of a modal operator.

Much of the intrigue with quantified modal logic concerns the interaction of quantifiers and modal operators. The simplest presentations of the standard S5 system include both the Barcan formula (\( \forall x \Box \phi \supset \Box \forall x \phi \)) and its converse (\( \Box \forall x \phi \supset \forall x \Box \phi \)) as theorems. Here I will show how an

\(^{\text{\footnotesize 23}}\)See Chapter 3, Section 4.
idealist semantics with the current, non-standard interpretation of the modal operators works for a standard system with the Barcan formulas as theorems. But let me say that I believe there are good reasons to deny the Barcan formulas, even within the context of the current semantics. But I have not yet figured how to reject the Barcan formulas within the semantics gracefully. I believe it can be done, but I will have to save this project for another day. For now, I must rest content with showing what would validate the Barcan formulas within the idealist semantics.

The Barcan formula and its converse are valid when (i) each individual idea in an ordered pair in the domain is part of the complete idea in that same ordered pair, and (ii) each complete idea in an ordered pair in the domain has the individual idea in that same ordered pair as a part. When the quantifier has the widest scope (e.g., in $\forall x \Box F x$), the formula says something about all the individuals in the actual world. So we may think of the quantifier in such a case as appearing within the scope of a veiled actuality operator. Within such a veiled actuality operator, the quantifier ranges over ideas of actual individuals—here, that will be complete ideas of actual individuals. $\forall x \Box F x$ will be true in a model when every complete idea $[u_c]$ such that $\langle [u_c], [u_i] \rangle \in D$ is such that the following holds of its mate $[u_i]$ in the domain: $[[F x] \mid [u_i]] \in W$. That is, $\forall x \Box F x$ is true when each complete idea $[u_c]$ is such that its corresponding individual idea $[u_i]$ is such that $W$ contains the propositional idea $[[F x] \mid [u_i]]$. If that’s the case, there’s no possibility of any actual thing not being F; everything actual is necessarily F.

When the necessity operator has widest scope (e.g., in $\Box \forall x F x$), the quantifier ranges over individual ideas. $\Box \forall x F x$ is true in a model when every individual idea $[u_i]$ such that $\langle [u_c], [u_i] \rangle \in D$ is also such that the propositional idea which consists in $[F x]$ being part of $[u_i]$ is itself part of $W$. That is, $\Box \forall x F x$ is true when each individual idea $[u_i]$ appearing in $D$ is such that $W$ contains the propositional idea $[[F x] \mid [u_i]]$. If that’s the case, there’s no possibility of something not being F; that is, necessarily, everything is F.

5.12 The Formalism

There are many ways to present formally the core semantic commitments from the previous section. The following presentation is not the most efficient, but it is among the clearest I’ve imagined.

Let’s begin with the primitive symbols of the language. They include:
individual constants (names) $a, b, ...$, with or without numerical subscripts

individual variables $x, y, ...$, with or without numerical subscripts

for each $n$ greater than 0, $n$-place predicates $F, G$, with or without numerical subscripts

symbols $\neg$ (negation), $\supset$ (conditional), $\forall$ (universal quantifier), $\Box$ (necessity operator), $(, \text{ and })$

I’ll also use various Greek letters, with or without subscripts, as metalinguistic variables, including $\alpha$ for terms—constants and variables—and $\Pi$ for predicates. I’ll use $\phi$ and $\psi$ for well-formed formulas (wffs), whose well-formedness is determined by the following formation rules:

(i) If $\Pi$ is an $n$-place predicate and $\alpha_1, ... \alpha_n$ are terms, then $\Pi \alpha_1, ... \alpha_n$ is a wff.

(ii) If $\phi$ is a wff then so are $\neg \phi$ and $\Box \phi$

(iii) If $\phi$ and $\psi$ are wffs and $\alpha$ is a variable, then $\neg \phi$, $(\phi \supset \psi)$, and $\forall \alpha \phi$ are wffs.

Without rehearsing them, I’ll adopt the standard definitions for $\lor$ (disjunction) and $\land$ (conjunction) in terms of $\neg$ and $\supset$, the definition of the existential quantifier $\exists$ as equivalent to $\neg \forall \neg$, and the definition of the possibility operator $\Diamond$ as equivalent to $\neg \Box \neg$. I’ll also take for granted the definition of a quantifier’s scope as well as the definitions of free and bound variables.

The interpretation function $I$ provides the denotations of names and $n$-ary predicates as they appear in wffs. The denotation of a name is an idea, either a complete idea (if the name doesn’t appear within the scope of a modal operator) or an individual idea (if the name does appear with within the scope of a modal operator). The denotation of an $n$-ary predicate within a wff is a set of ideas. $I$ obeys the following constraints:

(1) if $\alpha$ is a constant, then $\langle \{ \alpha_c, \alpha_i \} \rangle \in \mathcal{D}$ and $I(\alpha) = \{ \alpha_c \}$ if $\alpha$ does not occur within the scope of a modal operator and $I(\alpha) = \{ \alpha_i \}$ if $\alpha$ occurs within the scope of a modal operator.

(2) if $\Pi$ is an $n$-place predicate, then $I(\Pi \alpha_1, ... \alpha_n)$ is the set $\{ \{ \Pi \chi, \alpha_n \}, \{ \Pi \alpha_1, ... \chi \} \}$, which contains ideas represented by abstractions from every argument place in $\Pi \alpha_1, ... \alpha_n$. 

114
In Section 11, I introduced the notion of a variable assignment, but I did not formally define it. Let me provide that definition now:

\[ g \text{ is a variable assignment for } \langle D, I \rangle \text{ iff } g \text{ is a function that assigns to each instance of the variable } \alpha, \text{ some complete idea } [u_c] \text{ such that } \langle [u_c], [u_i] \rangle \in D \text{ if the instance of } \alpha \text{ occurs outside the scope of any modal operator, and the individual idea } [u_i] \text{ in that same } \langle [u_c], [u_i] \rangle \text{ if the instance of } \alpha \text{ occurs inside the scope of any modal operator.} \]

Variable assignments allow us to assign truth values to quantified formulas. But the addition of modal operators complicates the treatment of quantified formulas. There are three basic ways quantifiers and modal operators interact. The semantic treatment of each kind of interaction involves two steps.

Step 1. Does the universal quantifier itself occur within the scope of a modal operator? If no, then the universal quantifier \( \forall x \) ranges over complete ideas in the domain. Otherwise, the quantifier ranges over individual ideas in the domain.

Step 2. Do the instances of the variable bound by the universal quantifier \( \forall x \) occur within the scope of a modal operator? If no, then the admissible assignments to the instances of those variables include only the complete ideas in the domain. Otherwise, the admissible assignments to those instances of the variable include only the individual ideas in the domain.

Using these two steps, I will show how the semantics handles the three basic ways a quantifier and modal operator may interact.

**No Modal Operator.** A quantifier may not interact with any modal operators at all. Consider once more the non-modal formula \( \forall x Fx \). Since the universal quantifier does not occur within the scope of a modal operator, it ranges over complete ideas. And since ‘Fx’ also does not occur within the scope of a modal operator, the admissible assignments to \( x \) are limited to complete ideas. As a result of these two facts, I say that \( Fx \) is true no matter which complete idea in \( D \) \( x \) temporarily denotes, not just when \( x \) denotes the complete idea \( g \) assigns to this instance of \( x \). A variable assignment \( g^\alpha_{[u_c],[u_i]} \) (1) assigns \([u_c]\) to instances of \( \alpha \) outside the scope of any modal operator, (2) assigns \([u_i]\) to instances of \( \alpha \) within the scope of some modal operator, and (3) assigns the same
ideas to terms as $g$ with the possible exception of $[u_c]$ and $[u_i]$ to $\alpha$. We can then exploit the fact that, in the metalanguage, we can quantify over every ordered pair in the domain to say that `$Fx$' is true when $x$ denotes the complete idea of any ordered pair in $D$. The formula is true under a variable assignment $g$ just in case, for every ordered pair $([u_c], [u_i]) \in D$, `$Fx$' is true under the variable assignment $g^x_{[u_c],[u_i]}$. In this case, `$\forall x Fx$' is true when (Step 1) for every complete idea $[u_c]$ such that $([u_c], [u_i]) \in D$, (Step 2) `$Fx$' is true when $x$ temporarily denotes $[u_c]$.

**Quantifier First.** A quantifier may not occur within the scope of a modal operator while at the same time any variable it binds does occur within the scope of that modal operator. Consider `$\forall x \Box Fx$'. Since the quantifier does not occur within the scope of a modal operator, it ranges over the complete ideas in the domain. However, since `$Fx$' within that formula occurs within the scope of a modal operator, the admissible assignments to $x$ are limited to the individual ideas in the domain. Since `$x$' is bound by the quantifier which is not itself in the scope of a modal operator, the admissible assignments of ideas to $x$ are limited to the individual ideas which are paired up with complete ideas in the domain. `$\forall x \Box Fx$' is true when (Step 1) for each complete idea $[u_c]$ such that $([u_c], [u_i]) \in D$, (Step 2) the result of assigning its co-member $[u_i]$ to the instance of $x$ in `$Fx$' is true.

**Modal Operator First.** Finally, a quantifier and any variables it binds may occur within the scope of a modal operator. Consider `$\Box \forall x Fx$', for instance. Since the quantifier occurs within the scope of a modal operator, it ranges over individual ideas in the domain. And since `$Fx$' within that formula occurs within the scope of the modal operator, the admissible assignments to $x$ are limited to the individual ideas in the domain. `$\Box \forall x Fx$' is true when (Step 1) for each individual idea $[u_i]$ such that $([u_c], [u_i]) \in D$, (Step 2) the result of assigning the individual idea $[u_i]$ to the instance of $x$ in `$Fx$' is true.

In $D$, the complete idea $[u_c]$ in each ordered pair $([u_c], [u_i])$ teams up with exactly one different individual idea $[u_i]$, and the individual idea $[u_i]$ in each ordered pair $([u_c], [u_i])$ teams up with exactly one complete idea $[u_c]$. Given this 1-1 correspondence, the treatments in Quantifier First and Modal Operator First coincide with respect to Step 1. As a result, the formal treatment of quantification requires us only to take into account Step 2, whether the instances of any variable bound by the quantifier occur within the scope of a modal operator (and not whether the quantifier
itself appears within the scope of a modal operator). And that work is already done by the variable assignment.

We may now define the denotation of any term. Let $M$ be a model, $g$ be a variable assignment, and $\alpha$ be an instance of a term. We define $[\alpha]_{M,g}$, the denotation of $\alpha$ (relative to $M$ and the assignment $g$), as follows:

$$[\alpha]_{M,g} = \begin{cases} I(\alpha), & \text{if } \alpha \text{ is a constant} \\ g(\alpha), & \text{if } \alpha \text{ is a variable} \end{cases}$$

Relative to a model and variable assignment, the denotation of an instance of a constant is whatever idea $I$ assigns to it. Relative to a model and variable assignment, the denotation of an instance of a variable is whatever idea the variable assignment already assigns to the variable.

We may now use the interpretation function $I$ and the domain $D$ to introduce the notion of an intensional model for quantified modal logic. Let a model be an ordered quintuple $\langle A, W, P, D, I \rangle$ where $A$ is the idea of the actual world, $W$ is the idea of being a world, $P$ contains the conditions on ideational parthood (which I will discuss shortly), $D$ is the domain which houses pairs of complete and individual ideas, and $I$ is the interpretation function.

The valuation function defines the truth conditions of any wff under a model $M$ and a variable assignment $g$. Here is one way to define the valuation function for an intensional model:

(Vφ) For any $n$-place predicate $\Pi$ and any terms $\alpha_1, \ldots, \alpha_n$, $V_{M,g}(\Pi \alpha_1 \ldots \alpha_n) = 1$ if for each $m$, $1 \leq m \leq n$, the propositional idea $[[\Pi \alpha_1, \ldots, \alpha_{m+1}, \ldots, \alpha_n]_i [\alpha_m]_{M,g}]_i A$. $V_{M,g}(\Pi \alpha_1 \ldots \alpha_n) = 0$ otherwise.

(V¬) $V_{M,g}(\neg \phi) = 1$ iff $V_{M,g}(\phi) = 0$

(V⊃) $V_{M,g}(\phi \supset \psi) = 1$ iff either $V_{M,g}(\phi) = 0$ or $V_{M,g}(\psi) = 1$

(V∀) $V_{M,g}(\forall \alpha \phi) = 1$ iff for every $\langle [u_c], [u_i] \rangle \in D$, $V_{M,g}^{[u_c], [u_i]}(\phi) = 1$.

(V□) $V_{M,g}(\Box \phi) = 1$ iff $[\phi]_I W$.

(Vφ) deserves a handful of remarks. First, it generalizes the truth conditions of non-relational atomic formulas to any atomic formula. In the monadic case, (Vφ) implies that ‘Fa’ is true just in
case the propositional idea [Fa] is part of \( \mathcal{A} \). That is, ‘Fa’ is true just in case \([[[F x]_i [a, c]]_i \mathcal{A}] \). In the case of dyadic predicates, \((\forall \phi)\) implies that ‘Gab’ is true just in case the propositional ideas \([[[G x b]_i [a, c]]_i \mathcal{A}]\) and \([[[G a x]_i [b, c]]_i \mathcal{A}]\) are parts of \( \mathcal{A} \). And so on.

Second, I will assume that the actual world is harmonious (See Chapter 3). Harmonious models are such that for any wff \( \Pi \alpha_1 \ldots \alpha_n \), a propositional idea \([[[\Pi \alpha_1, \ldots, x, \alpha_{m+1}, \ldots, \alpha_n]_i [\alpha_m]]_i \mathcal{A}]\) for some \( m \) is itself part of \( \mathcal{A} \) if and only if for every \( m, 1 \leq m \leq n \), the proposition idea \([[[\Pi \alpha_1, \ldots, x, \alpha_{m+1}, \ldots, \alpha_n]_i [\alpha_m]]_i \mathcal{A}]\) is part of \( \mathcal{A} \). So, for example, the propositional idea of Solomon being David’s son is part of \( \mathcal{A} \) if and only if the the propositional idea David being Solomon’s father is part of \( \mathcal{A} \).

The auxiliary assumptions in the previous section justify what remains of the valuation function. Besides \((\mathcal{A} \neg)\), \( \mathcal{A} \)'s completeness and consistency help ensure that \((\mathcal{V} \neg)\) provides the correct truth conditions for negated formulas. \((\mathcal{A} \supset)\) helps ensure that \((\mathcal{V} \supset)\) provides the correct truth conditions for conditionals. The remarks on quantification in the previous section help ensure that \((\mathcal{V} \forall)\) provides the correct truth conditions for universally quantified formulas, whether the quantifier appears within or without the scope of some modal operator. Finally, \((\mathcal{W} \supset), (\mathcal{W} \neg)\), and the remarks on names and quantification in Sections 8 and 11 help ensure that \((\mathcal{V} \Box)\) provides the correct truth conditions for formulas modal operators.

This section has touched on each member of the quintuple \(\langle \mathcal{A}, \mathcal{W}, \mathcal{D}, I, P \rangle\) except for \( P \), which contains principles for which ideas are parts of which other ideas. In the next section, I survey these principles and explain why the current interpretation of the formalism, along with the axioms of the mereology for ideational parthood, at least partially justify the axioms and theorems of S5.

### 5.13 Parthood

One could abstract away from my interpretations of \( \mathcal{A} \) as the idea of alpha and \( \mathcal{W} \) as the idea of being a world to a purely formal semantics to deny that \( \mathcal{W} \) is part of \( \mathcal{A} \) or that ideational parthood is transitive.\(^{24}\) One could also opt for another interpretation of \( \mathcal{W} \) as part of an applied semantics for deontic logic, for example, and then deny that \( \mathcal{W} \) is part of \( \mathcal{A} \) or that ideational parthood is transitive. But my interpretation of the semantics is part and parcel of offering it as a way to model metaphysical

\(^{24}\)I show how the pure formalism works in validating the axioms of different modal systems in Chapter 2, Section 4.
reality, which is arguably captured best by the S5 system. It just so happens that if we adopt the principles for ideational parthood I survey in Section 4, we have good reason to accept (K), (T), (S4), (B4), and (S5)—all theorems or axioms of S5—as true.

With (N), \((W \supset (W \land))\) and the remarks on quantification help secure the necessity of the theorems of first-order logic. For example, given \((W \land)\), if \([\phi \land \psi]\) is part of \(W\), \([\phi]\) is, too. Then, based on our intuitive understanding of ‘if... then...’, \([(\phi \land \psi) \supset \phi]\) is also part of \(W\). We then infer from (N) that \((\phi \land \psi) \supset \phi\) is necessarily true. Intuitively, all other theorems of first-order logic have corresponding propositional ideas as parts of \(W\), making those theorems necessarily true. These considerations justify an important inference rule in the weakest normal modal system K, a system whose theorems include all the theorems of first-order logic:

**Necessitation Rule.** If \(\phi\) is a theorem of K, so is \(\Box \phi\).

Additionally, \((W \supset)\) secures K’s characteristic axiom schema:

\[(K) \Box (\phi \supset \psi) \supset (\Box \phi \supset \Box \psi)\]

Let’s try to falsify (K) by making the antecedent true and the consequent false. The antecedent \(\Box (\phi \supset \psi)\) is true just in case \([\phi \supset \psi]\) is part of \(W\). The consequent \(\Box \phi \supset \Box \psi\) is false when \(\Box \phi\) is true (i.e., when \([\phi]\) is part of \(W\)) but \(\Box \psi\) is false (i.e., when \([\psi]\) is not part of \(W\)). Given \((W \supset)\), we cannot falsify (K). To do so, we’d need to have \([\phi \supset \psi]\) and \([\phi]\) but not \([\psi]\) as parts of \(W\). But \((W \supset)\) says that if \([\phi \supset \psi]\) and \([\phi]\) are parts of \(W\), so is \([\psi]\). So (K) is true in models which respect \((W \supset)\).

For (T) and (S4) especially, it will be useful to implement some more shorthand. Consider the atomic proposition ‘Fa’. According to the previous section, ‘Fa’ is true when the propositional idea 
\([[\[F \geq x\] \vDash a_c]]\) is part of \(A\), and it is necessarily true when 
\([[\[F \geq x\] \vDash a_i]]\) is part of \(W\). Thus, we may no longer use ‘[Fa]’ as the propositional idea of a’s being F without specifying whether the idea of a involved is its complete or individual idea. When an atomic formula \(\phi\) appears within the scope of \(\Box\), let \([\phi]\) represent the propositional idea which is part of \(W\) when \(\Box \phi\) is true. When an atomic formula \(\phi\) does not appear within the scope of \(\Box\), let \([\phi]\) represent the propositional idea which is part of \(A\) when \(\phi\) is true under a variable assignment.

\[(T) \Box \phi \supset \phi\]. The idea of being a world in general is part of the idea of being this world. Therefore, if ideational parthood is transitive, then any propositional idea which is part of \(W\) is also
part of \( A \). Suppose \( [\phi] \) is part of \( W \), which means that \( \phi \) is necessarily true. Then, given transitivity, \( [\phi] \) is part of \( A \). Given that \( [\phi] \) is part of \( W \), any ideas of individuals involved in \( [\phi] \) are individual ideas. So if a non-negated propositional idea \( [[Fx]; a_i] \) is involved in \( [\phi] \), transitivity implies that if \( [\phi] \) is part of \( A \), the idea \( [[Fx]; a_i] \) is also involved in \( [\phi] \) as part of \( A \). But now the first of the I-C principles kicks in: if \( [[Fx]; a_i] \) is involved in \( [\phi] \) as part of \( A \), then \( A \) has as a part the same propositional idea which involves \( [[Fx]; a_i] \) instead of \( [[Fx]; a_i] \). Similar remarks apply to \( \phi \) when it involves negation, given the second I-C principle. If, for example, \( [\neg Fa] \) is part of \( W \), then the idea which is part of \( W \) is \( [[Fx] \not< a_i] \). But given transitivity, if \( [[Fx] \not< a_i] \) is part of \( W \) and \( W \) is part of \( A \), then \( [[Fx] \not< a_i] \) is part of \( A \). But, given the second I-C principle, if \( [[Fx] \not< a_i] \) is part of \( A \), so is \( [[Fx] \not< a_i] \), which covers the truth of ‘\( \neg Fa \)’. Given the I-C principles and the transitivity of parthood, what is necessarily true is true and every instance of (T) is true.

(In most of what follows, I will gloss over whether a propositional idea involves an individual or complete idea of an individual. This will simplify the presentation considerably.)

**S4** \( \Box \phi \supset \Box \Box \phi \). If \( \Box \phi \) is true, then the idea \( \phi \) is part of \( W \), the idea of being a world. \( W \) is a special kind of idea. Given that \( W \) includes “what it takes” to be a world in general, we might wonder whether part of what it takes to be a world is that there is an idea of being a world in general which has ideas as parts. \( W \) plays a crucial role in determining what is necessary and possible. Unless we think that there might have been no modal truths—that it isn’t part of \( W \) that \( W \) exists—then \( W \) must somehow implicate itself as a necessary being. But if \( W \) implicates itself as a necessary being—if the idea that \( W \) exists is itself part of \( W \)—then we have good reason to suppose that \( W \)’s having various parts is also part of being a world in general. Could an idea have had different parts and been the same idea it is? In the case of ideas, it isn’t implausible to think that different content makes for a difference in identity. So if \( [\phi] \) is part of \( W \), which means that \( \phi \) is necessary, we have good reason to suppose that having \( [\phi] \) as a part is itself part of \( W \), which means that \( [\phi] \)’s being part of \( W \) is necessary, which means, finally, that \( \phi \) is necessarily necessary. On this line of reasoning, every instance of S4 is true.

**B**. \( \phi \supset \Box \Box \phi \). Since \( W \) is part of \( A \), both are consistent if \( A \) is. Now suppose \( \phi \) is true, that \( [\phi] \) is part of \( A \). Then \( A \)’s consistency guarantees that \( [\neg \phi] \) isn’t part of \( W \). But it also seems reasonable that the preconditions for worldhood could not have included that \( \phi \) is false. That is, it is a precondition of worldhood that there is no such precondition. So not only is \( [\neg \phi] \) not part of \( W \)

---

120
(due to \(W\)’s being part of \(A\) and \(A\)’s consistency), but the idea that \(\neg \phi\) isn’t part of \(W\) is itself part of \(W\). Given this line of reasoning, if \(\phi\) is part of \(A\), it’s part of \(W\) that \(\neg \phi\) isn’t part of \(W\). This restriction on \(P\) fortifies the truth as necessarily possible and so secures the truth of (B).

(S5). \(\Diamond \phi \supset \Box \Diamond \phi\). In the discussion about (S4), I claimed we had good reason to suppose that \(W\)’s having the very parts it has is itself part of \(W\). If that’s true, then we might wonder whether \(W\)’s not having the ideas as parts that it doesn’t have as parts is also part of \(W\). If we think \(W\) implicates itself as a necessary being and that a difference in parts would make for a different idea, we have good reason to suppose that if \(\neg \phi\) isn’t part of \(W\), then it is part of \(W\) that \(\neg \phi\) isn’t part of \(W\). If there are necessary truths about which ideas \(W\) doesn’t have as parts, then \(W\) has parts that concern which properties it doesn’t have as parts. Given this restriction in \(P\) concerning \(W\)’s parts, (S5) is true.

The restrictions on \(P\) which would make (S4), (B), and (S5) true are not themselves axioms of the property mereology. They are instead plausible principles about \(A\) and \(W\) specifically. But given the general picture of ideational parthood and what kinds of ideas \(A\) and \(W\) are, I find these principles fairly reasonable.

5.14 Malebranche

Thus far I’ve developed the semantics under the assumption that the ideas involved belong to God’s mind. In this section, I sketch a brief history of theories of divine illumination leading up to Malebranche’s Vision in God theory. In the next section, I abstract a structure from that theory as a model for the metaphysics and epistemology of modality and then explain what resources it has for resolving some difficulties in modal epistemology.

From the neo-platonists onwards, there is a tradition according to which God’s ideas also function as our own raw materials for thought. The metaphor often used in connection with God’s sharing his ideas is one of illumination. Drawing from Plato and the Neo-platonists, Augustine compares God’s illumination of our minds with the sun’s illumination of the visible world. He identifies the eternal truths with ideas in God’s mind, and claims we know them because he illumines our minds with them.\(^{25}\) Working in the wake of Augustine, both Leibniz and Malebranche hold that God illumines

\(^{25}\)For references, see Chapters 3 and 4 in Menn (2002).
our minds with some eternal truths. Prior to creation, the Leibnizian God has ideas of all possible individuals, which He concatenates in various ways to produce complex ideas of an infinite number of possible worlds. With all possible worlds before His understanding, God then actualizes the best one among them. For Leibniz, God’s understanding is the “realm of eternal truths,” which includes modal truths. He is also “the source of what reality there is among possibilities.” Not only does the necessity or possibility of a truth depend on the contents and relations among ideas in the divine understanding, the very being of necessary and possible truths depends on the divine understanding.

Compared to Leibniz, Malebranche has less to say about the divine illumination of modal truths. But Malebranche’s Vision in God theory says more about the mechanics of illumination itself. Malebranche alters and extends the scope of more traditional illumination theories. Whereas Augustine and Leibniz hold that the objects of illumination are truths, Malebranche argues that the objects of illumination are the ideas involved in those truths. For instance, whereas Augustine might have said that God illumines our minds with the truth that 2 and 3 equals 5, Malebranche claims that God illumines our minds with the ideas which are 2, 3, and 5 and that the equality of 2 and 3 to 5 is a relation among those ideas.

Malebranche also extends the scope of illumination to cover not only the knowledge of what’s necessary and eternal, but both thought and perception generally. The Vision in God theory claims that the mental lives of human beings are wholly dependent on God. When we think of a triangle, for instance, we think of the idea of intelligible extension in a specific way. When we see a triangular object, we sense the idea of intelligible extension in a specific way. In both cases, the idea of the triangle is an idea in God’s mind. The Vision in God theory is relevant here because of its characterization of our ideas as God’s own. So I will restrict my attention to Malebranche’s account of ideas as the shared raw materials of human and divine thought.

The Vision in God theory says that created minds enjoy a close union with the divine mind—so close, in fact, that our minds rely on the same materials for thought. Malebranche appeals to this closeness to explain how knowledge of created beings is possible for us:

---

28 Malebranche (1991, 3.2.6)
...God must have within Himself the ideas the ideas of all the beings He has created (since otherwise He could not have created them), and thus He sees all these beings by considering the perfections He contains to which they are related. We should know, furthermore, that through His presence God is in close union with our minds, such that He might be said to be the place of minds as space is, in a sense, the space of bodies. Given these two things, the mind surely can see what in God represents created beings, since what in God represents created beings is very spiritual, intelligible, and present to the mind. Thus, the mind can see Gods works in Him, provided that God wills to reveal to it what in Him represents them. (230)

We can be sure that God, as creator, has ideas of whatever he creates. Without an idea of something, God couldn’t intend to create it. Ideas, then, are archetypes of created beings. 29 God’s ideas fully capture every created thing, and the content of each idea derives from God’s considering his own perfections. Given that God’s ideas completely and accurately capture every created being, we can think of created beings through thinking God’s ideas of those creatures. If God’s perfect and complete ideas are our own, how does Malebranche explain ignorance and error? Our grasp of God’s ideas is incomplete and imperfect, but that is a function of our limited cognitive powers, not the ideas themselves which our limited powers grasp. 30 Thus, Malebranche writes:

As the ideas of things in God include all their properties, whoever sees their ideas can also see all their properties successively; for when we see things as they are in God, we always see them in perfect fashion, and the way we see them would be infinitely perfect if the mind seeing them were infinite. What is lacking to our knowledge of extension, fires, and motion is the shortcoming not of the idea representing it but of our mind considering it. (237)

29.“God contains within Himself in an intelligible fashion the perfection of all the beings He has created or can create, and that through these intelligible perfections He knows the essence of all things, as through His volitions He knows their existence.” (Elucidation X, in Malebranche (1991, 617).

30.“The things the mind perceives through illumination or through a clear idea it perceives in very perfect fashion, and it even sees clearly that whatever obscurity or imperfection there is in its knowledge is due to its own weakness and limitation or some lack of attentiveness on its part, and not to the imperfection of the idea it perceives.” (Elucidation X in Malebranche (1991, 621).
God’s sharing his ideas with us explains the knowledge of created objects we do have, imperfect and incomplete though it is. Thus, God’s having an idea of \( x \) explains how \textit{someone} could know something about \( x \). God’s sharing his idea of \( x \) with us explains how \textit{we} could know something about \( x \).

Let’s consider the example of our grasping a triangle in Euclidean geometry. God has a perfect idea of a Euclidean triangle, in general—let’s call this an idea of triangularity. God shares this idea of triangularity with us. When we think of triangularity or any triangles in particular, the material of this thought is God’s own idea of triangularity. If ideas were the unique property of those who grasped them, then Tom would have his idea of triangularity, you would have yours, and I would have mine. Then, given that we incompletely and perhaps sometimes inaccurately represent what a triangle is, each of our ideas would have different content. Such a view threatens the objectivity of geometry. But our incomplete and imperfect grasp of triangularity does not threaten the objectivity of geometry in Malebranche’s theory. The number of distinct incomplete and imperfect thoughts of Euclidean triangularity is not explained by an equal number of incomplete and imperfect ideas of Euclidean triangularity. There is only one such idea, God’s, and he shares it with us, though we each may grasp it in different and more or less complete ways. Thought, for Malebranche, resembles perception in the following way: you and I may both perceive Fido, but we do so in virtue of perceiving different aspects of Fido. Similarly, you and I grasp a single thing, God’s idea of triangularity, but we may do so in virtue of perceiving different aspects of that idea.

God is omnipotent and, in Malebranche’s occasionalist system, the only substance which acts by its own efficacy.\textsuperscript{31} God’s ideas are extensions of his divine efficacy in a metaphorical but real sense. Unlike most contemporary accounts of the inhabitants of a “third realm,” Malebranchean ideas are causally efficacious:

\begin{itemize}
  \item It is certain that ideas are efficacious, since they act upon the mind and enlighten it, and since they make it happy or unhappy through the pleasant or unpleasant perceptions by which they affect it. Now nothing can act upon the mind unless it is superior to it—nothing but God alone; for only the Author of our being can change its modifications.
  \item All our ideas, therefore, must be located in the efficacious substance of the Divinity,
\end{itemize}

\textsuperscript{31} Malebranche (1991, 236).
which alone is intelligible or capable of enlightening us, because it alone can affect intelligences. (232)

God’s ideas actively affect the understanding, and the understanding has a corresponding, “purely passive” faculty.\textsuperscript{32} But how should we understand the notion of an efficacious idea in relation to Malebranche’s occasionalism?\textsuperscript{33} First, Malebranche distinguishes genuine from occasional causes but, in general, both kinds of causes are events or states of affairs.\textsuperscript{34} The substance responsible for a genuine cause has an efficacy of its own—its efficacy is one of its own dispositional features. But no occasional cause features a substance with an efficacy of its own. An occasional cause does not act by its own efficacy. Instead, an occasional cause meets a certain condition under which, in general, God efficaciously wills something else to happen. Occasional causes channel God’s efficacy. In the divine understanding are general volitions, divine willings for how things should go given certain conditions. The laws of nature, for example, are themselves some of the general volitions (or the intentional content of some general volitions). When one ball hits another, the event is an occasional cause, which, given the divine general volitions governing collisions, results in their different velocities. The divine general volitions are efficacious, and God’s omnipotence is responsible for their efficacy. Similar remarks hold for psychological events of the understanding. Acts of the understanding are the result of previous (or perhaps simultaneous) occasional causes along with divine general volitions governing human cognition. These general volitions channel God’s efficacy through God’s own ideas, which are the objects of our understanding and what our passive faculties receive in cognition.

Malebranche also places limits on the knowledge we have of causal relations. According to Robert Adams, Malebranche claims to see that the efficacy of divine volitions follows from the infinite perfection of the divine being even though Malebranche has “no distinct idea of how the effects follow from the divine volitions.”\textsuperscript{35} In his \textit{Christian and Metaphysical Meditations}, Malebranche presents Christ as saying,

\begin{quote}
I give human beings no distinct idea that corresponds to the word power or efficacy […]
\end{quote}

For even if you believe that God does [or makes, fait] what he wills, it is not that you

\textsuperscript{32} Malebranche (1991, 236).
\textsuperscript{33} The summary here of Malebranche’s views on God’s efficacy rely heavily on Adams (2013).
\textsuperscript{34} Adams (2013, 104).
\textsuperscript{35} Adams (2013, 92).
see clearly that there is a necessary connection between the will of God and the effects, since you do not even know what the will of God is. But what is evident is that God would not be omnipotent if his absolute volitions remained inefficacious.36 (MCM IX.2: OCM X.96)

According to Malebranche, we can know that God’s volitions are efficacious but we cannot know what that efficacy consists in. We can know that God’s volitions cause various things, but we do not know how God does so. We may apply these positive and negative knowledge claims to the Vision in God theory itself. Malebranche would claim, I believe, that the ideas we consider as ours are some of God’s. But he would also claim that we have no distinct idea of how God shares these ideas with us or how they are causally efficacious. This ignorance is inescapable not so much because of the natures of the relata (God’s ideas and our minds) but because of the nature of the relation itself. We have no more an idea of how God’s ideas cause us to understand than we do of how one billiard ball collides with another to change its velocity. Both require an occasional cause which meets some condition specified in a law (a divine general volition), but how exactly God channels his efficacy into the effect is a mystery.

5.15 Divine Idealism

At the beginning of the chapter, I argued that there’s room for a modal semantics whose meanings are mathematically simple enough that human beings could know that a proposition is necessarily true without first knowing that the proposition is analytic or the negation of a contradiction. I also argued that there’s room for a modal semantics whose meanings on some intended interpretation are the sort of objects to which we have cognitive access. In this section, I sketch a brief theory of modal idealism and explain whether and to what extent it helps explain the possibility of modal knowledge.

The view that abstract objects such as propositions are ideas or concepts in God’s mind has received some attention more recently. Adams (1983, 749–750) argues that such a view could explain how, given a causal connection between divine and human minds, we could have the kind of modal knowledge that has no bearing on survival. Under the label “theistic activism,” Morris and Menzel (1986) defend a kind of modal idealism against a modal variant of the Euthyphro dilemma. Although

36Quoted from Adams (2013, 86).
Leibniz places possible worlds in the divine understanding; he does not employ possible worlds to define necessity or possibility in the usual way. Contemporary theistic views of modality go one step further than Leibniz and characterize necessity and possibility using possible worlds. Both Pruss (2011) and Leftow (2012) argue that God’s power at least partially grounds modal reality and then associate possible worlds with divine ideas or concepts. But if we define necessary truth as what is true in every possible world, identifying possible worlds with divine ideas or concepts will not help make sense of how finite creatures like us could grasp those necessary truths. The view below appeals to the formalism from this chapter to help explain how specifically finite minds could have modal knowledge. Then, like Adams, the view posits a causal connection between divine and human minds to help explain how we have any connection to the modal truths we seem to grasp.

Let me begin with a brief description of who, or what, I take God to be. Here, the meaning of ‘God’ is not necessarily tied to a specific religious tradition, but it isn’t intended to be anti-religious either. What’s left when we strip off the specifically Christian content found in Leibniz and Malebranche is a picture of omniscient mind responsible for the being of all things, including ideas. These ideas are the raw materials of God’s own thought. But he’s not responsible for them in a Cartesian way: ideas are not dependent on God’s will. He couldn’t have decided that $2 + 2 = 5$ or that 2 would not exist. Ideas do not depend on God’s will for their existence, but on his understanding. God continuously holds them in existence by understanding them. Understanding a proposition is more like entertaining than believing a proposition. God understands false propositions but doesn’t believe them or think they’re true.

Let’s suppose, like Malebranche, that God’s ideas also function as the raw materials of our own thought. There is a common store of ideas in God’s mind, and human thought consists in grasping them. But we do not usually grasp them perfectly or completely. Consider, for example, God’s idea of something so simple as a quark. Physicists have some idea of the functional role of the quark, but no one knows everything there is to know about the nature of quarks. It is doubtful whether anyone could know everything there is to know about quarks. What’s true for something so simple as a quark is true of things much more complicated, such as the biological property of life, individual persons, and far off planets. Our ideas of these things are not only incomplete. For reasons of hubris, misleading information, and so on, our ideas of these things also include ideas which are not included in the divine ideas of those things. Since modal truths arise from the relations between divine ideas,
our incomplete ideas mean that we are often in the dark about which modal propositions are true. And our imperfect ideas mean that we are sometimes wrong about what is possible or necessary. But insofar as we have good evidence that our ideas correctly represent things as they are, we have good evidence for thinking that we represent things as they could or must be.

Consider a modal truth, say, that mammals are animals necessarily or essentially. How might we know this? We do not know this by looking at all the possible ways mammals are across an infinite number of possible worlds. Nor do we have a complete grasp of the idea of being a mammal, which may be, for all we know, infinitely complex. Instead, we have some incomplete grasp of being mammalian, which includes the idea of being an animal. The parts of the idea of being mammalian determine what mammals are essentially, and, in turn, what is possible for mammals. How do we know that being an animal is part of being mammalian? We use something like Husserl’s method of clarification. We first think of being a mammal in general. We then entertain an idea similar to the first which either includes an additional idea or lacks an idea. We think, for example, of an idea which includes the idea of not being an animal. We may then discover that the resulting idea is not identical to what we first thought. Or we may discover that the resulting idea includes two ideas, one of which is the negation of the other (e.g., being an animal and being not an animal). In either case, we may discover that our first idea of being mammalian did in fact include the idea of being an animal. Thereafter, when we imagine what is possible for mammals, we do so with the awareness that it includes the idea of being an animal. We needn’t run through an infinite number of possibilities to see whether they all possible mammals are animals. Nor do we need a perfect and complete knowledge of the idea of being mammalian and the idea of being an animal.

Let us also suppose, like Malebranche, that ideas are causally efficacious. This supposition seems to fit descriptions of our cognitive phenomenology. Ideas are entertained and understood. In virtue of their content, there is something it is like to entertain and understand them. Ideas are experienced. We describe them sometimes as being hidden from us, other times as discovering them, and yet other times as being struck by them. They can inspire within us a sense of euphoria, and we sometimes get lost in them. We grasp them, and there is an experience of grasping them. Husserl writes: “among the different obscure movements of thought which stir us, one thought, for example,
stands out from all the rest and has a sensitive effect on the ego, as it, so to speak, forces itself against the ego.”³⁷ These descriptions of our understanding suggest that ideas are causal agents.

There is a passage in Frege which suggests the same. In his essay, “The Thought,” Frege says that thoughts are not usually the sorts of things we would call “actual” (wirklich). For ” the world of actuality is a world in which this acts (wirkt) on that and changes it and again undergoes reactions (Gegenwirkungen) itself and is changed by them.”³⁸ But soon after, he applies this causal language to thoughts themselves:

How does a thought act? By being grasped and taken to be true. This is a process in the inner world of a thinker which may also encroach on the sphere of the will and make itself noticeable in the outer world as well. ... When a thought is apprehended, it at first only brings about changes in the inner world of the apprehender, yet it remains untouched in its true essence, since the changes it undergoes involve only inessential properties. There is lacking here something we observe throughout the order of nature: reciprocal action. Thoughts are by no means unreal but their reality is of quite a different kind from that of things. And their effect is brought about by an act of the thinker without which they would be ineffective, at least as far as we can see. And yet the thinker does not create them but must take them as they are. They can be true without being apprehended by a thinker and are not wholly unreal even then, at least if they could be apprehended and by this means be brought into operation.³⁹

Fregean thoughts are propositional. They are what one grasps when one thinks. And here Fregean describes them as effective. One may contend that such descriptions should ultimately concern mental acts and not the ideas involved in those mental acts. But if there are mental acts, what is acted upon? Is there not something grasped? Does this not still suggest that ideas have, at the very least, a passive power of being grasped? Unlike Malebranche, then, we needn’t suppose that our faculty of understanding is purely passive and the objects of cognition purely active in their interaction. If one

³⁷ Husserl (1939, 77).
³⁸ Frege (1956, 343).
side is more active than the other, we needn’t say which one. What is important to the current view is that there is a causal relation between minds and the ideas they grasp.\(^{40}\)

To my knowledge, no account of abstract objects in recent decades claims for them any kind of causal efficacy. No one has suggested that the solution to variants of the Benacerraf problem—that we have no knowledge abstract objects due to their causal inefficacy—is to deny the premise that abstract objects are causally inefficacious. I suspect this is so because attributing causal powers to mind-independent abstract objects in this way would elevate them to the status of substances. But there is little to no reason to elevate the raw materials of thought to substances. Doing so would appear to be an ad hoc solution to the Benacerraf problem. But if the raw materials of thought are mind-dependent, their causally efficacy is more easily explained. On the current view, they are mind-dependent objects which channel the efficacy of a mind which holds them in being. There are enough raw materials and they have enough efficacy because, according to the current view, they belong to an infinite, omnipotent being. The causal efficacy of ideas is a natural extension of the view that the raw materials of thought ultimately belong to such a being. Those with much more naturalistic tendencies may think such a view is unlikely to be true. But, if true, it would partly dispel the mystery of our epistemic access to so-called abstract objects. Like perception, what explains our epistemic access to some object is a causal connection.

But one may reasonably wonder whether this story has any more explanatory power than the traditional theory that abstract objects are acausal. For the theory leaves open how God shares his ideas with us in two ways. First, it is noncommittal about the metaphysics of the divine mind as well as human minds. Perhaps Spinozism is true. Perhaps human minds are Leibnizian monads. Or perhaps some form of dualism is true of human minds or mental properties. Or perhaps human minds are wholly material (though I doubt this position can be meaningfully explained or understood). Second, the view is noncommittal about the nature of causation. It only requires that the nature of the divine mind and the natures of human minds are not so fundamentally different to preclude their

\(^{40}\)If there is a causal relation between minds and the ideas they grasp, then ideas have causal powers. And if ideas have causal powers, then either there is something that can be done to ideas or something ideas can do. And these powers will not best be understood as reducible to inclusion and exclusion relationships between ideas. But the current view is not one according to which modality is ultimately reducible to relationships between ideas, but a view about how we represent and modal facts and about what modal propositions are.
causal interaction. Depending on one’s philosophical temperaments, one may count these explanatory lacunae against or in favor of the theological structure abstracted from Malebranche’s theory.

I would contend that the nature of causality, like the metaphysical nature of substance, is ultimately inaccessible to us. It is a virtue of a theory which leaves mystery where mystery ultimately belongs, given our cognitive and conceptual limitations. It is a theoretical vice to require of a theory what no theory can ultimately provide. Even though we do not know what causation is or what substances are, we know, in some sense of ‘know’, that there is such a thing as causation and such things as substances which causally interact. If we are in the business of assigning pluses and minuses to theories for their explanatory power, then a kind of Malebrancheanism receives a plus both for its ability to explain that a causal relation accounts for our access to the objects by which we have knowledge of modal reality and for its refusal to explain what cannot be explained anyway—the nature of causation itself.
CHAPTER 6

Set-theoretical Idealism

6.1 Introduction

A set is a mathematical object which contains its members. There are sets of numbers, sets of people, and sets of sets. There are infinitely large sets, like the set of the natural numbers, and there is an empty set with no members whatsoever. This is only a small sampling of what we seem to know about set-theoretical reality. What could sets be, metaphysically speaking, so that we could know what we seem to know?

One view is that sets are acausal abstract objects with no location in space or time and immune to any possible pushes or pulls. But there is a two-way highway between us and the set-theoretical universe, and this acausal view of sets renders traffic in each direction quite mysterious. On the way up, we know about sets and therefore have some sort of access to them. But absent a causal connection to sets, what other kind of connection could we have to set-theoretical reality which would explain how we know about its inhabitants? On the way down, some sets contain concrete objects like you and me. But how could abstract objects contain concrete objects? If containment isn’t a spatio-temporal relation, what kind of relation could it be? There are two mysteries, then: our knowledge of set-theoretical truths and the meaning of the truths we claim to know.

Another view is that each set is “nothing over or above” its members, that the existence of a set is nothing but or little more than the co-existence of its members. This view would explain how the two-way highway between us and set-theoretical reality is possible. On the way up, knowledge of a set would reduce to knowledge of its members. Knowing about a set of numbers would reduce to knowing about the numbers; knowing about a set of people would reduce to knowing about the people. And on the way down, if a set’s existence consisted merely in the co-existence of its
members, we would have a good idea of what set-theoretical containment is as well as what various set-theoretical truths mean. Sets would contain numbers, people, and sets in the same way, namely, through the co-existence of their members. And the meaning of a truth about a set’s members would concern the coexistence of the set’s members.¹

Yet this second view is also problematic. First, there would be no empty set. With no members, there is nothing for a set’s existence to consist in. Second, any single-membered sets would be identical to their single members. The existence of each single-membered set would consist merely in the existence of its single member. Yet it is now standard to assume that the empty set exists and that any singleton is distinct from its sole member.² This second view implies that we don’t know some of what we seem to know about set-theoretical reality.

These two traditional views of the metaphysical theories of sets are problematic. The first renders the meaning of set-theoretical truths and our knowledge of those truths equally mysterious. The second denies that we know some of what we seem to know. So what could sets be so that we could know what we seem to know? This question is unavoidable, if we seek a philosophical understanding of set theory.³ I argue that if sets are ideas, we could know what we seem to know about set-theoretical reality.

### 6.2 Preliminaries

Widespread views about sets together yield various puzzles. The goal of the present chapter is to provide a metaphysical theory of sets which justifies those views but also resolves the puzzles to which they give rise. Other metaphysical theories of sets may do the same, of course. I merely want to clear out a space on the table for my preferred theory to show that it resolves these puzzles in interesting and, in my view, plausible ways.

I take standard views about the content and practices of contemporary set theory for granted because mathematics seems to be an autonomous discipline. Mathematicians do not, and need not,

¹If existence is univocal, that is.
²For instance, the Axiom of Regularity (also called the Axiom of Foundation) in Zermelo-Frankel set theory rules out sets which are members of themselves. But if a singleton is identical to its lone member, then we may infer that the singleton is a member of itself via the indiscernibility of identicals.
³Parsons (1983, 503) claims that he “sees no way to obtain philosophical understanding of set theory while avoiding metaphysics…”
ask permission from scientists or philosophers to posit the existence of a mathematical object. At least with respect to philosophy, David Lewis captures this sentiment exactly:

Renouncing [sets] means rejecting mathematics. That will not do. Mathematics is an established, going concern. Philosophy is as shaky as can be. To reject mathematics for philosophical reasons would be absurd. If we philosophers are sorely puzzled by the [sets] that constitute mathematical reality, that’s our problem. We shouldn’t expect mathematics to go away to make our life easier. Even if we reject mathematics gently—explaining how it can be a most useful fiction...—we still reject it, and that’s still absurd... How would you like the job of telling the mathematicians that they must change their ways...?  

Although contemporary set theory gives rise to various puzzles about the metaphysics and epistemology of sets, mathematicians needn’t put their work on hold while philosophers develop theories of sets. We let mathematicians tell us which mathematical objects exist and what mathematical properties they have. Some philosophers reject this attitude and deny the existence of some or all mathematical objects.  

Nothing I say here will change their minds. In fact, they are welcome to use my proposal as part of a *reductio* against the existence of sets. They may grant that my proposal offers the most plausible resolution to various longstanding puzzles about the metaphysics and epistemology of sets and so argue that sets would be the sorts of things I say there are, if they existed at all. Since, as they may argue, my proposal is obviously ridiculous, we have all the more reason to reject sets altogether. The importance of what follows consists in the explanatory value of my proposal, not in whether I’ve definitively nailed down the metaphysical essence of sets.

While we may accept what mathematics tells us about which mathematical objects exist and what mathematical properties they have, we need not assume that mathematics tells us everything about mathematical objects. There are meaningful questions about the nature of mathematical objects which mathematical theories themselves leave open. For example, are sets human ideas? Set theory doesn’t say explicitly whether they are or not. But the question is apparently meaningful and would seem to have a right answer. Presumably, sets are not human ideas: there are not enough human

---


5 Field (1980) and Hellman (1989).
ideas to account for the set-theoretical hierarchy. Human beings are finite in number collectively and
finite in time individually. So the number of human ideas would seem to be finite, too. But there are
infinitely many sets. Since presumably sets outnumber human ideas, sets are not human ideas. This
explanation for why sets are not human ideas doesn’t immediately fall out of set theory. It’s not as if
set theory contains a theorem which says, “No set is a human idea.” Rather, we point to a deficiency
in the metaphysical theory to account for a set-theoretical truth about the number of sets. But if sets
are not human ideas, what are they? This question is also apparently meaningful and set theory does
not answer it either. We may judge a metaphysical theory of sets not only for how poorly it accounts
for set-theoretical truths, but also how well it accounts for them.

I doubt we’ll ever know what sets are. Such is the nature of metaphysical theorizing, in my view.
But popular views of what sets are conflict with one another and cannot all be true. And some of
those views conflict with or undercut standard mathematical practice. What’s more, human beings
seem to have this peculiar ability to grasp set-theoretical truths, but it isn’t clear at all how we’re able
to grasp those truths or what kinds of things those truths concern. Here is the question: what could
sets be such that creatures like us are able to grasp truths about them? The challenge is to provide
a metaphysical theory of sets which not only resolves various puzzles but is also sensitive to the
question of how we’re able to grasp set-theoretical truths.

Before I describe these puzzles, it will be beneficial to introduce the widely shared views about
sets which give rise to them.

6.3 Membership

Sets have members. We may express the idea that \( x \) is a member of a set \( A \) by saying that \( x \) belongs
to \( A \) or that \( A \) contains \( x \). Or we may use a formal language to express the same idea. In such a
language, \( 'x \in A' \) means that \( x \) is a member of \( A \). ‘A’ is a name for a set, but the name itself provides
no clue about which things, if any, it contains. So we may also use names for sets which make their
members explicit. Such a name consists of a pair of brackets enclosing a list or description of the
set’s members. For instance, we write \( '\{\text{Sally, Tom}\}' \) to name the set whose members are Sally and
Tom, if there is such a set.
Sally and Tom are not sets but people, and if a set contains either one them, there is a set with some non-set member. We’ll call any sets which have some non-set as a member *impure* and any sets which do not have any non-sets as members *pure*. It is common for set-theorists not only to discuss impure sets but to point to them as paradigmatic cases of sets. Here are a few examples:

A pack of wolves, a bunch of grapes, or a flock of pigeons are all examples of sets of things. .... Sets, as they are usually conceived, have *elements* or *members*. An element of set may be a wolf, a grape, or a pigeon.⁶

Examples of finite sets are: the set of the inhabitants of a city, the set of hydrogen atoms in the sun, and the set of natural numbers from 1 to 1000...⁷

A set (or class) is a collection of objects. These objects may be numbers, functions, physical objects, or even sets. Since there are no restrictions on the objects which may be members of sets...⁸

Although mathematicians often point to impure sets as paradigmatic cases of sets, they just as often restrict themselves to the pure sets, since impure sets are unnecessary for defining the natural, rational, and real numbers or examining, say, the Axiom of Choice. For example, soon after Halmos and Schoenfield use impure sets to introduce their topic, they restrict themselves to the world of pure sets:

It is important to know that a set itself may also be an element of some other set. ... What may be surprising is not so much that sets may occur as elements, but that for mathematical purposes no other elements need ever be considered.⁹

It turns out that [pure sets] are sufficient for mathematical purposes; and they are also sufficient to illustrate all the problems which arise in the general case. We shall therefore restrict ourselves to this case, and henceforth take *set* or *class* to mean *pure set*.¹⁰

---

⁶Halmos (1974, 1).
⁸Shoenfield (1967, 238).
⁹Halmos (1974, 1).
¹⁰Shoenfield (1967, 238).
These and similar pronouncements don’t imply that there are no impure sets but only that there is a range of mathematical topics whose full understanding requires only pure sets.\textsuperscript{11} Charles Parsons writes:

The literature on the foundations of set theory does not sufficiently emphasize that the exclusion of individuals in the standard axiomatizations of set theory is a rather artificial step, taken for the convenience of pure mathematics.\textsuperscript{12}

On the whole, authors seem to assume that impure sets exist and that the pure versus impure distinction is not merely conceptual.

Why should we think there are impure sets? Mathematics applies to the world. We count and measure worldly objects and represent them with mathematical structures. Sets play an important role in grouping things, computing, and counting, and set-theoretical structures within model theory help us model and represent worldly structures. But if there are no impure sets, set theory would seem to be inapplicable to anything outside of set theory. On the importance of using impure sets to count, Potter writes:

It is by no means obvious what justifies the applicability of mathematics in general to what lies outside it, and it may well be that the reduction of mathematics to set theory does not supply such a justification. But even if set theory’s role as a foundation of mathematics turned out to be wholly illusory, it would earn its keep through the calculus it provides for counting infinite sets. The most natural, if not literally the only, way to ensure that the calculus is available to be applied to counting non-mathematical things — chairs, electrons, thoughts, angels — is to allow such things into the theory as individuals.\textsuperscript{13}

In the middle of a case for impure sets, Timothy Smiley and and Alex Oliver provide some examples of counting which seem to require impure sets:

\textsuperscript{11} Against Oliver and Smiley (2006, 146), who claim that these authors are confused about mathematical purposes and mathematical objects. Within the scope of ‘mathematical purposes’, Oliver and Smiley want to include things like the application of mathematics to the natural world. But the contexts of these passages make clear that the authors have in mind the purposes of pure mathematics.

\textsuperscript{12} Parsons (1983, 504, n. 2).

\textsuperscript{13} Potter (2004, 51).
...‘If there are n people, how many pairs of people are there?’ Given that this simple question is legitimate (substitute cards or horses and think of its importance for poker players or bookmakers), further questions must be legitimate too. How many pairs of pairs? How many pairs of pairs of pairs? We are led inexorably to full-blown applied set theory.\textsuperscript{14}

Mathematical practice attests to the incredible usefulness of impure sets. And without impure sets, “the applicability of set theory has to be given up.”\textsuperscript{15} We may restrict ourselves to the pure sets as mathematicians often do, if our goal is to explain in the simplest terms possible the axioms which govern the set-theoretic universe. But if we are concerned about what sets are, metaphysically speaking, we ought not restrict ourselves to the pure sets. For if we restrict ourselves to pure sets, we risk confusing some merely accidental feature common to pure sets for an essential feature of sets in general. Given that there are both pure and impure sets, we may judge a metaphysical theory of sets for how well it accounts for the ability of sets to “contain” both sets and non-sets. Is the set containing Tom a set in the same sense as the set containing the empty set? If so, does the set which contains Tom contain Tom in the same sense that the set which contains the empty set contains the empty set? What about the set which contains both Tom and the empty set? A metaphysics of sets which captures contemporary set theory will explain set theory’s applicability.

\section*{6.4 Desiderata}

Set theory itself defines neither the notion of a set nor the notion of set membership. Although set theory does not explicitly define the notions of set or set-membership, its axioms implicitly define them. Zermelo-Fraenkel set theory with the Axiom of Choice (ZFC) has become the standard system for axiomatic set theory. Given the metaphysical conservatism from Section 1, ZFC’s axioms place two kinds of constraints on what to expect from a metaphysical theory of sets. First, a metaphysical theory of sets should not conflict with ZFC’s axioms. Second, a metaphysical theory of sets has more or less explanatory power given how well it explains why sets have the sorts of features the axioms

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{14}Oliver and Smiley (2006, 146).
\item \textsuperscript{15}Zermelo (1996, 1232).
\end{itemize}
\end{footnotesize}
attribute to them. Below, I pinpoint some set-theoretical facts which fall out of the axioms and which any conservative metaphysics of sets should explain.\(^\text{16}\)

Let us first nail down some conventions for discussing set-theoretical matters. First, let lowercase letters near the end of the alphabet (e.g., \(x, y, z\)) be variables ranging over both sets and also whatever non-sets which are eligible for set-membership. Second, let upper-case letters near the end of the alphabet (e.g., \(X, Y, Z\)) function as variables ranging over sets only. Then, let \(\in\) and \(=\) signify the set-membership and identity relations, respectively.

The first feature a conservative metaphysics should explain is the extensionality of sets, which falls out immediately from the Axiom of Extensionality. Extensionality says that any set \(X\) and any set \(Y\) are identical if and only if they have the same members and implies that no two sets have the same members. That is, \(X\) and \(Y\) are the same set when every member of \(X\) is a member of \(Y\), and every member of \(Y\) is also a member of \(X\). It is widely held that a theory without Extensionality is not a theory of sets.\(^\text{17}\) George Boolos (1971, 23), for example, hesitates to claim that Extensionality is analytic for broadly Quinean worries. But he nevertheless believes that it is “far more central” to the concepts of set and set-membership than the other axioms. Any metaphysical account of sets must explain why no two sets have the same members.

ZFC’s axioms also imply that there is a unique set with no members at all. Depending on the presentation, some axiom or axioms secure the existence of an empty set. And extensionality implies that such a set is unique. Some axiomatic presentations secure the existence of the empty set with its own axiom.\(^\text{18}\) Other presentations imply that the empty set exists through more roundabout means. If, by whatever means, we already know that there is a set, then we may use the Separation Schema to infer the empty set’s existence. The schema says that for every predicate \(\phi\), the statement ‘for every set \(A\), there is a set \(B\) containing all and only the members \(x\) of \(A\) such that \(\phi(x)\)’ is an axiom. That is, once we have a set \(A\), we may define any condition \(\phi\) and form a set \(B\) of all the members of \(A\) which satisfy \(\phi\). This schema is technically not an axiom though every one of its instances is an axiom. Different instances of the schema specify different predicates in the place of \(\phi\). When every member of set \(X\) is a member of set \(Y\), \(X\) is a subset of \(Y\). Given any set, Separation ensures

\(^{16}\)For a complete list of ZFC’s axioms, see Jech (2003, 3-13).

\(^{17}\)Potter (2004, 33).

\(^{18}\)See Enderton (1977, 18), for example.
the existence of every one of its subsets. Since the empty set has no members at all, it is vacuously true that every one of its members is a member of any set. Given any set, the empty set is one of its subsets. To infer the empty set’s existence from Separation, however, we first need a set for which we can pinpoint a condition \( \phi \) which none of its members satisfy. If we already know of such a set, and I would contend that we do—we know of all kinds of sets—we can infer that the empty set exists. But if we limit our set-theoretical claims to what we may infer from the axioms, we cannot yet assume that there are any sets, let alone an empty set. 19

In those presentations without an empty set axiom, the axiom which secures the empty set uses another axiom to do so. The latter is the Axiom of Union, which says that for any set \( A \), there is a set \( B \) whose members are all and only the members of \( A \)’s members. Union allows us to crack open any set and dump its members’ members into one set. By Extensionality, the union of a set is unique. Given the set \( \{\{x, y\}, \{w, z\}\} \), for instance, its union is \( \{x, y, w, z\} \). The Axiom of Infinity relies on the Axiom of Union and ensures that there is an empty set. Its formulation requires us to define two special set-theoretic properties. First, call a set \( B \) a successor of set \( A \) just in case \( B \) is the union of \( A \) and \( \{A\} \). By the Union Axiom, then, \( A \)’s successor has \( A \) as a member and all of \( A \)’s members as members. Second, call a set \( A \) inductive just in case, first, the empty set \( \emptyset \) is a member of \( A \), and, secondly, if any set is a member of \( A \), then its successor is a member of \( A \). The Axiom of Infinity says that an inductive set exists. By definition, then, an inductive set exists only if the empty set does, too. The empty set comes along for free with the existence of an inductive set.

There are other justifications of the empty set, of course. Penelope Maddy (1997: 43) claims that the empty set’s existence follows from the iterative conception of set. David Lewis (1991: 12) claims we “better believe in it with the utmost confidence” because it plays a crucial role in set theory’s role as a foundation for mathematics. Whatever justification we give for the empty set, it is now standard to assume that it exists. To say that the empty set exists is to say that there exists a set which has no members—it is a set, it is something, but it contains nothing. Since it is now standard to believe the empty set exists, a conservative metaphysics of sets will explain how it is something even though it contains nothing.

Set-theorists also generally accept that there are singleton sets which are distinct from their lone members. Most textbooks, like Jech (2003: 7), introduce singletons as a special application of the Axiom of Pairing. According to this axiom, for any \( x \) and \( y \), there is a set \( \{x, y\} \) whose members are exactly \( x \) and \( y \). Given any \( x \) and \( y \), pairing guarantees the existence of a set whose members are exactly \( x \) and \( y \). Pairing and Extensionality together guarantee that there is a unique set whose members are exactly \( x \) and \( y \). If \( x \) and \( y \) happen to be the very same individual, then \( \{x, y\} = \{x\} \), a singleton set with exactly one member.

Another axiom secures the distinctness of each singleton from its lone member. The Axiom of Regularity, which I will not produce here, precludes the existence of three kinds of sets.\(^{20}\) No set is self-membered. No set exhibits a membership loop. This means no set is among its members, its members’ members, etc., all the way down through the minimal elements ultimately required to form that set. And no set exhibits an infinitely descending chain of membership: no set has some member, which has a member, which has a member, and so on, without end. It is the first prohibition which concerns singletons. If a singleton is identical to its lone member, we may infer that the singleton is a member of itself via the indiscernibility of identicals. But Regularity implies that no set is self-membered. Hence, no singleton is identical to its member. A conservative metaphysics of sets will explain what sets are such that singletons are distinct from their sole members.

Set theory as standardly practiced has three more features in need of explanation. There is only one ‘\( \in \)’ symbol in set theory. So unless the membership symbol is ambiguous, which would be surprising, we should assume that ‘\( \in \)’ is univocal and signifies in every instance a single kind of membership relation. In light of the applicability of set theory to the natural world, the univocality of ‘\( \in \)’ implies that both pure and impure sets have members in the same sense. Set-theoretical containment is obviously a metaphor. But what kinds of things are sets that they can contain, in the same sense, something with no spatio-temporal location and something with a spatio-temporal location? How can we understand that metaphor in a way that would respect standard mathematical practice?

The iterative conception of sets is now widespread. In fact, Michael Potter (2004: 52) claims “the idea that the only coherent conception is the iterative one has become widespread.”\(^{21}\) The iterative


\(^{21}\)Besides Potter (2004, Ch. 3), see Boolos (1971).
conception characterizes the set-theoretical universe as a hierarchy of stages. Depending on one’s purposes, the first stage consists either of individuals or nothing at all. Each successive stage in the hierarchy consists of all and only the sets which can be formed from whatever is in the previous stage. For example, if Tom and Harry alone occur in the first stage, then \{Tom\}, \{Harry\}, \{Tom, Harry\} and \emptyset occur in the second stage. No set occurs at any stage of the hierarchy without each of its members occurring somewhere previously in the hierarchy. The iterative conception precludes the existence of sets which would falsify the Axiom of Regularity. That is, it precludes non-well-founded sets.

Why should we endorse the iterative conception (and also endorse the Axiom of Regularity)? Simon Hewitt (2015, 21) considers two possible justifications. The first, which Hewitt endorses as “prudent,” says that the iterative conception is simply analytic of the set concept. There may be set-like entities which are non-well-founded. But they are not sets. They are not sets for the simple reason that they conflict with the iterative conception. The second justification, which Hewitt tables, is that sets depend for their existence on their members, but not the other way around. Sets appear in the hierarchy after their members on the iterative conception because stages encode the metaphysical dependence of sets on their members. But Hewitt worries whether the appeal to metaphysical dependence secures the iterative conception and well-foundedness, in particular. Do we have good reason to believe that the relation of metaphysical dependence is itself well-founded? If metaphysical dependence is not itself well-founded, we cannot appeal to it to justify the iterative conception and explain why we should reject non-well-founded sets.

Whether or not the iterative conception is analytic of the set concept, one who provides a metaphysics of sets should wonder whether the account explains why there are no non-well-founded sets. Given that the non-existence of non-well-founded sets is non-contingent, what about the nature of sets explains why they are well-founded? I suspect that the most plausible answers will appeal to some notion of metaphysical dependence, not necessarily the notion of metaphysical dependence—I am skeptical of such a thing—but some metaphysical relation which requires a specific kind of dependence of one relatum on the other. What could that relation be? How do sets depend on their members?

Finally, a conservative metaphysics of sets has more explanatory power insofar as it explains why there are set-theorists with standard practices and beliefs in the first place. The challenge is
to provide a resolution to a Benacerraf-style worry about set-theoretical knowledge. Benacerraf argues that the most natural semantics for arithmetical statements treats a statement like ‘1 is odd’ the way we normally treat a statement like ‘Fred is tall.’ Naturally, ‘1 is odd’ is true just in case ‘1’ denotes 1, presumably an abstract object, which is also odd. Benacerraf argues, however, that the best epistemology says that our knowledge of the world depends on our causal interaction with it. Abstract objects aren’t spatial or temporal, so we don’t have any causal interaction with them. Therefore, we’re stuck with either a non-standard semantics or a substandard epistemology. Even though the causal theory of knowledge is no longer en vogue, Benacerraf’s dilemma raises an important question: if numbers are abstract, how can we know anything about them? By similar reasoning about sets, we attain a similar difficulty. We seem to know quite a bit about sets. Can we provide a metaphysics of sets which explains our set-theoretical knowledge?

In summary, the set-theoretical hierarchy has a handful of features which those who seek to offer a conservative metaphysics of sets should explain. We may judge such an account against how well it answers these questions:

(i) Why are there no two sets with the same members?

(ii) What is set-membership, and how do sets “contain” both spatio-temporal and non-spatio-temporal entities?

(iii) What is the empty set and why should we believe there is one?

(iv) Why are singleton sets distinct from their lone members?

(v) What explains the well-foundedness of sets? If sets depend on their members metaphysically, how should we understand that dependence?

(vi) What kind of epistemic access do we have to sets?

Over the next two sections, I consider and reject two general kinds of metaphysical accounts of sets. Both accounts flounder on one or more of questions (i) through (vi). Then I present and evaluate my set-theoretical idealism.

6.5 Highbrow Sets

Expositions of set-theory typically include one of two informal views about sets. Often, single authors use both. There is, on the one hand, what Max Black calls “the highbrow view” of sets as something like abstract objects which contain their members in some figurative sense. Many have tried to flesh out this figure of speech with more figures of speech, likening these abstract containers to physical containers such as sacks or lassos. Some authors stick with containers. Enderton (1977, 3), for example, contrasts the empty set with the singleton set of the empty set when he says “a man with an empty container is better off than a man with nothing—at least he has the container.” The highbrow view has some explanatory power. But that explanatory power is weak if not totally illusory.

The highbrow view initially seems to plausibly answer questions (iii) and (iv) and thus able to explain what the empty set is and why singleton sets are distinct from their members. The highbrow view says that the empty set is a sort of abstract object, perhaps an empty abstract container. Such a view would seem to explain why the empty set is something rather than nothing. It would also seem to explain why singleton sets are distinct from their lone members. No container of whatever kind contains itself, not if the container is a sort of boundary which contains only whatever is inside that boundary. So if there are abstract containers which contain a single thing, we have good reason to believe that the container is distinct from the contained.

The problem is that containment is a spatio-temporal concept. So set-theoretical containment is metaphorical. How should we understand the metaphor, then? Without a clearer understanding, explanations which bank on the metaphor are written checks for an empty account. These explanations ultimately ring hollow in light of the highbrow view’s failure to explain something deeper, namely the issue in question (ii). That question asks what set-membership is and how sets “contain” both spatio-temporal and non-spatio-temporal entities. If sets are abstract objects of any kind, it is difficult to see how they could relate to both abstract and non-abstract objects in the same way. After all, sets do not contain their members by covering them with ethereal slime. Furthermore, containment

---

23 Boolos (1971) attributes the lasso metaphor to Kripke. Dedekind likened sets to sacks during H. Felix Bernstein’s 1897 visit to Harzburg. (See Bernstein (1996)). Some authors stick with containers, for example, Enderton (1977, 3), who, when contrasting the empty set with the singleton set whose lone member is the empty set, says “a man with an empty container is better off than a man with nothing—at least he has the container.”

24 Potter (2004: 22), for example.
occurs not only not from a distance, but from no distance at all. But sets may bear a temporal relation to their members. Unless individuals exist eternally, as soon as they come into or go out of existence, so do sets which contain them, or sets which contain sets which contain them (and so on). Until we have some deeper explanation of what set-theoretical membership is on a highbrow view, we cannot appeal to it to explain otherwise puzzling features of sets.

Whatever explanatory power the highbrow view has is swamped by its apparent inability to answer questions (i), (ii), and (v). Having already briefly discussed (ii), let’s move on to (i): why do sets obey the Axiom of Extensionality? Surely Extensionality itself does not exert some kind of gate-keeping force which denies existence to any set which would share all the same members with another. And surely there are not infinitely many possible sets we simply stipulate out of existence upon adopting Extensionality as an axiom. And surely Extensionality isn’t written indirectly into every single member of every single set, as if each thing’s nature included infinitely many stipulations of the form that it and \( x_1 \), and \( x_2 \), and ... , and \( x_n \) are members of only one set. The correct answer, in outline, is that sets are, by nature, the kinds of things which satisfy the Axiom of Extensionality.

But what specifically is in the nature of a set which precludes the existence of another with the same members? Identifying sets with abstract objects doesn’t alone guarantee that there aren’t two or more sets with the same members. And I can’t think of any more specific account of abstract objects which could explain why none of them, as sets, have the same members. One may speculate as a last resort that the truth of Extensionality is some sort of brute fact in need of no explanation at all. But in that case, we should count ourselves very fortunate to have grasped a brute fact about abstract objects which have no location in space or time and no causal powers to interact with our cognitive apparatus. In any case, the highbrow view of sets doesn’t seem to permit an answer to question (i) about why there are no sets with the same members.

There is also the longstanding problem of how we know anything about abstract objects if they do not partake in the causal order.\(^{25}\) This particular problem doesn’t presuppose that knowledge requires a causal connection, even though the problem, as Benacerraf originally put it, does presuppose a causal theory of knowledge. Even if we put aside the causal theory of knowledge, we should remember why it would have seemed so intuitive to someone.\(^{26}\) We can trace a causal connection

\(^{25}\) Benacerraf (1973).

\(^{26}\) Goldman (1967).
through perception and testimony to many of the things about which we know something. And how we trace that causal connection illuminates how we’ve come to know what we know. How do I know that Fido is brown? I saw him. How do I know that the universe is many billions years old? I’ve read and heard many different kinds of testimony. But now suppose I claim to know something. And I also claim to lack any causal connection to what that knowledge concerns. The question naturally arises, then: how do I know what I claim to know, if I am not causally connected to whatever my knowledge concerns? What other connection might there be? Identifying sets with abstract containers seems to cut off any possible explanation for our set-theoretical knowledge, and, hence, any hope to answer question (v) about our epistemic access to sets.

6.6 Lowbrow Sets

There is what Black calls “the lowbrow view” of sets as pluralities, where a plurality of objects is nothing more than the objects themselves.27 The lowbrow view says that the set \{Tom, Dick, and Harry\} is no more than Tom, Dick, and Harry. The lowbrow view says that sets are pluralities. That position is distinct from the view which says that all pluralities are sets. This second position answers the question of when a plurality forms a set, namely always. The question is similar to the Special Composition Question about when some things compose an object.28 One may adopt the first and reject the second position by requiring that only some pluralities meet some further condition to secure their status as a set. Whatever that condition might be and whether it would conflict with some standard view or practice within mathematics does not concern me here. For the trouble with the lowbrow view is that it alone conflicts with now standard views in set theory. Below, I describe some motivation for the lowbrow view and then explain why it conflicts with standard views in set theory.

If the highbrow view seems unlikely to overcome the problems surveyed in the previous section, perhaps the solution is to rid sets of their containers and identify them with the pluralities of their members. Ditching abstract objects in favor of pluralities has some immediate payoffs, some of

---

27 Black (1971, 616) calls these pluralities “collections.” But that label is now also used for sets, even when the speaker has an abstract container view in mind. Hewitt, (2015: 5) argues that Black is a set-theoretical nihilist, that there are no sets. This gloss on Black’s position had never occurred to me because his arguments about the highbrow and lowbrow views of sets occur concern what sets are not whether they are.

which I mentioned at the beginning of the paper. First, the lowbrow view provides a compelling answer to question (i) about why are there no two sets with the same members. Whereas it is difficult to explain why there aren’t two indiscernable abstract containers which contain the same individuals, it is not difficult to explain why there not two or more pluralities which consist of the same individuals. A plurality of individuals is nothing more than the individuals. Same individuals, same plurality. Hence, the only way for two pluralities to be distinct from one another is for one to have an individual the other lacks. So if sets were pluralities, we would have the kind of explanation we desired for the truth of the Axiom of Extensionality.

Identifying sets with pluralities would also de-mystify the containment metaphor and explain how sets contain both spatio-temporal objects and abstract objects as members in the same sense of “contain”. If a set is a plurality, then a set contains an individual just in case an individual is among the individuals of the plurality. Individuals of vastly different kinds may form a plurality. There are pluralities of people and pluralities of numbers, but also mixed pluralities of people and numbers. It is equally mysterious what numbers are and what people are. But that Todd is among Todd, Julie, and the number 5 is not mysterious. More generally, whatever the natures of the individuals in a plurality are, being one among those individuals is not mysterious.

The lowbrow view also helps explain our epistemic access to some of set-theoretical reality. On the lowbrow view, knowledge of a set is no more mysterious than knowledge of its members. For if a set is the plurality of its members, knowledge of a set requires only that we know of its members. We have supposed that there are impure sets whose members include concrete individuals. We causally interact with some of these individuals. A causal connection to Fred (say, via perception) and a causal connection to Julie (also via perception) explains how we know of Fred and Julie. So our causal connections to Fred and Julie explains our awareness of Fred and Julie, the individuals which are together a plurality and also a set. The lowbrow view purports to explain our knowledge of some impure sets by appealing to the causal connection we have to their members. But the plurality view does not explain how we have any knowledge of the hierarchy of pure sets. Pure sets do not have any concrete individuals in their transitive closure. So unless pure sets also have causal powers, we will

---

29 Maddy (1980, 1990) develops a similar view of impure sets.
have no causal connections to members of any pure sets. This may not be as problematic as it seems. For as I will argue shortly, it’s doubtful that there could be any pure sets at all on the lowbrow view.

The lowbrow view accounts for the truth of Extensionality, how sets contain such disparate kinds of objects in the same sense, and our knowledge of some portion of set-theoretical reality. But the lowbrow view conflicts with standard set theory. Many of the early set-theorists and their contemporaries recognized that the lowbrow view renders the empty set suspect. Cantor says that a point-set which contains no points does not exist, strictly speaking. Frege claims that “if we burn down all the trees of a wood, we thereby burn down the wood. Thus there can be no empty class.” Zermelo says the the empty set is “not properly speaking (uneigentlich) a set.” Gödel suggests the empty set is fictitious. This skepticism about the empty set is justified if sets are pluralities. For a plurality of nothing is *nothing*, not something. As Black says, “collecting nothing at all to produce a unique object . . . is mystification on stilts.” There is no reason to believe in the empty set on the lowbrow view. And so it provides no resources for explaining what the empty set is, if we take contemporary mathematics at face value and assume there is one.

The lowbrow view also cannot explain why there are singletons which are distinct from their members. A plurality of a single individual, if that makes sense to say, is nothing more than *that individual*. But if any singleton is identical with its member, then, by the indiscernability of identicals, what’s true of one is true of the other. So from the identifiy of a set with its member, we may infer that the singleton is a member of itself, which violates the Axiom of Regularity and therefore conflicts with the iterative conception of set. Further complications arise in the case of singleton sets whose members are sets with two or more members. Orthodox set theory distinguishes \{Tom, Julie\}, whose members are Tom and Julie, from \{\{Tom, Julie\}\}, the singleton set whose lone member is \{Tom, Julie\}. If a singleton, like any set on the lowbrow view, is nothing but its members, then we shall have to say that \{\{Tom, Julie\}\} is \{Tom, Julie\}. And then we shall have to say that the set which is

---

30 See Oliver and Smiley (2013, 248) for the translation of and reference to Cantor’s remarks about the empty set.  
31 See Oliver and Smiley (2006, 126) for the translation of and reference to Frege’s remark.  
32 Zermelo (1967, 202).  
33 Gödel (1944, 131).  
34 Black (1971, 621–2).  
35 See Bolzano (1851, Sec. 3) and Moore (1962, 14).
\{\{\text{Tom, Julie}\}\} = \{\text{Tom, Julie}\} \text{ has both one member and two members.}^{36} \text{ To ward off contradiction, the lowbrow view should deny the existence of singletons altogether.}

Thus, a commitment to the lowbrow view has led some to formulate alternative axiomatic set theories with neither singleton sets nor the empty set.\(^37\) The lowbrow view has led others to reinterpret mathematical expressions about the problematic sets as a piece of convenient and ultimately disposable technical machinery.\(^38\) We need not survey these alternative theories or interpretations here. For they conflict with our current goal of developing a conservative metaphysics of sets which takes standard set theory at face value. Such a metaphysical account of sets grants that there is an empty set, that it contains nothing, and that it isn’t a convenient fiction. It will also grant the existence of singleton sets, that such sets are distinct from their members, and that they, too, are not convenient fictions. In the next section, I provide such an account and show how it answers questions (i) through (vi).

6.7 Set-Theoretic Idealism

What unites any number of individuals into a single set? Cantor himself did not think this was the work of an abstract container. Nor did he seem to think the plurality of individuals was itself a set so that any additional uniting was unnecessary. Instead, he seems to have attributed a set’s unity to thought: “by a ‘manifold’ or ‘set’ I understand in general any many (\textit{Viele}) which can be thought of as one (\textit{Eines}).”\(^39\) Cantor appears to say that sets are thoughts or ideas of their members, and that “‘collection’ (\textit{Zusammenfassung}) is an operation of the mind.”\(^40\) Most have found the characterization objectionable for that very reason.

Despite these protests, the idea has reappeared occasionally in more mathematical contexts. Rudy Rucker (1982, 40), for example, calls sets “thought balloons.” William SchAAF very explicitly says that sets are mental:

---

\(^{36}\)This is the argument in Frege (1893, Section 10, n. 17).
\(^{39}\)See Hallett (1986, 33) for a translation of and reference for this remark.
\(^{40}\)Parsons (1983, 506).
A set is the mental construct obtained by regarding several discrete things as constituting a single whole. Forming a set is thus a mental act: the human mind arbitrarily brings together certain things and regards the collection itself as a new kind of thing. This new thing is an artificial entity, in the sense that the unity lies entirely in the concept and not in the things themselves.\(^41\)

In his elucidation of the iterative conception of set, Hao Wang similarly suggests that sets are the result of tying individuals together in intuition and that infinite sets are the products of an idealized infinite intuition ‘overviewing an infinite range of objects’.\(^42\) Why should mathematicians as philosophically well-informed as Cantor and Wang appeal to the mind to elucidate the set concept? And why should mental language appear even in more informal elucidations of the concept? The answer, I suspect, is that set-theoretical idealism offers a compelling explanation of how many individuals unite to form a single set. But the explanatory power of set-theoretical idealism goes well beyond this. In this section, I develop a kind of set-theoretical idealism. In the next, I show how set-theoretical idealism affords compelling answers to questions (i) through (vi).

Sets are ideas. What are ideas, then? And which ideas are they? I’ll tackle the first question first. When I think of Obama, for example, there are three things to distinguish. There is the token mental act, the intentional object (Obama, in this case), and the intentional or representational content in virtue of which the mental act is about Obama. That content is an idea of Obama. And more generally, ideas are the intentional or representational contents which direct mental acts towards their objects.

Sets are a specific kind of idea. They are collective ideas of individuals. But there are many different kinds of ideas of individuals. Let me first explain what I mean by ‘idea of an individual’. Then I’ll explain the notion of a collective idea. In chapter 4, I distinguish between complete and individual ideas. A complete idea of an individual contains all the ideas which the individual satisfies. The complete idea of Obama, for instance, includes that he is the 44th U.S. President, married to Michelle, and the father of two girls, for instance. I characterize but do not define the individual idea of an individual as the idea which contains only those ideas which an individual must satisfy, no

\(^{41}\) Schaaf (1960, 11).

\(^{42}\) Wang (1983, 531)
matter how that individual could have been. If Obama is essentially a person, essentially a human
being, and so on, then his individual idea includes the ideas of being a person and being human. But
neither complete nor individual ideas are the ideas of individuals which comprise the collective ideas
in the present account of sets.

Take any actual individual. There is a spectrum of ideas of that individual ranging from
the most complete, its complete idea (which includes everything truly predicable of the individual), to
its individual idea (which includes only the ideas which the individual must satisfy), to a minimal
idea of the individual, the idea of that, where that is the individual in question. For lack of a better
name, I will call these minimal ideas haecceities. My notion of a haecceity corresponds to Husserl’s
notion of a determinable X, what in an intentional act is “the ‘‘bearer of predicates” and “the pure X
in abstraction from all predicates.” I’ll prefix a ‘=’ to an italicized name to signify its haecceity
(e.g., Tom’s haecceity is =Tom).

These haecceities have analogues in both the philosophy of language and metaphysics. They
are the mental analogues of Millian names or demonstratives which refer to an individual directly,
without the aid of any descriptive content. Like Millian names and demonstratives, they, too, lack
any descriptive content. Unsurprisingly, they are the mental analogues of the metaphysical notion of
what Robert Adam’s calls a thing’s thisness, a property of being a particular individual (or of being
identical to a particular individual). But whereas a thing’s thisness is the property of being that
particular thing, a haecceity, in my sense, is the idea of that thing (without any qualitative content).
They are, in a sense, the mental bare particulars which are sufficient to determine the intentional
object of an idea and on to which we attach other ideas to think of the object as being such-and-such.

Sets are collective ideas of haecceities. What makes an idea collective? There is in thought a
mental analogue of plural reference. We may use ‘Tom, Dick, and Harry’ to refer to Tom, Dick, and
Harry collectively. And we may have the idea Tom and Dick and Harry to think of Tom, Dick, and
Harry collectively. The occurrences of ‘and’ in ‘the idea Tom and Dick and Harry’ do not signify the
connective ‘and’ which conjoins propositions, sentences, or formulas of a language. Neither ‘Tom’
nor ‘Dick’ nor ‘Harry’ is or signifies a proposition, sentence, or formula. Nor does it signify the
phenomena of mental conjunction (from Chapter 4). If it did signify mental conjunction, and the

43 Husserl (1983, Section 131).
44 Adams (1979).
ideas of Tom, Dick, and Harry were to include ideas about their different heights, then the resulting mental conjunction would be the idea of a single thing which has three different heights. An idea of Tom, Dick, and Harry is not an idea of one thing whose features include some or all of Tom’s features, some or all of Dick’s features, and some or all of Harry’s features. Ordinarily, an idea of Tom, Dick, and Harry is an idea of three different persons, each with their own distinguishing characteristics. Call the ideas which one collects into a collective idea that collective idea’s components.

Set-theoretical idealism is the view that sets are collective ideas of haecceities. Let us run through some definitions for simple set-theoretical notions to see how set-theoretical idealism ultimately accounts for them.

**Membership.** Something x is a member of set A just in case xs haecceity is among the haecceities which comprise A. Consider Tom, Dick, and Harry, for example. Each has a unique haecceity: $=Tom$, $=Dick$, and $=Harry$. The set $\{\text{Tom, Dick, and Harry}\}$ is the collective idea $=Tom and =Dick and =Harry$, and that collective idea is an idea of Tom, Dick, and Harry. An individual $x$ is a member of $\{\text{Tom, Dick, and Harry}\}$ just in case xs haecceity is either $=Tom$ or $=Dick$, and $=Harry$. So $\{\text{Tom, Dick, and Harry}\}$’s members include Tom, Dick, and Harry. The set is an idea of its members.

**Subset.** Set A is a subset of set B just in case every one of A’s component haecceities is also a component haecceity of set B. For example, suppose A = $\{\text{Tom, Harry}\}$ and B = $\{\text{Tom, Dick, and Harry}\}$. According to set-theoretical idealism $\{\text{Tom, Harry}\}$ is the collective idea $=Tom and =Dick and =Harry$, and that collective idea is an idea of Tom, Dick, and Harry. An individual $x$ is a member of $\{\text{Tom, Dick, and Harry}\}$ just in case xs haecceity is either $=Tom or =Dick$, and $=Harry$. So $\{\text{Tom, Dick, and Harry}\}$’s members include Tom, Dick, and Harry. The set is an idea of its members.

Set-theoretical idealism also importantly captures the difference between membership and subset. First, being a subset of a set isn’t sufficient for being one of that set’s members. In the previous example, A is a subset of B, but it is not also a member of B. B’s members are Tom, Dick, and Harry and the collective idea $=Tom and =Dick$ is not itself Tom, Dick, or Harry. Conversely, being a member of a set isn’t sufficient for being one of that set’s subsets, even if that member is a set itself. For consider the set A = $\{\text{Tom and } \{\text{Dick and Harry}\}\}$ which is the collective idea $=Tom and =\{\text{Dick and Harry}\}$. The second component idea, $=\{\text{Dick and Harry}\}$, is the haecceity of the set B = $\{\text{Dick and Harry}\}$. B is a member of A since its haecceity is among A’s component haecceities. But B is
not also one of A’s subsets. B isn’t a subset of A because not all of its members are members of A. The haecceities of B’s members Dick and Harry (=Dick and =Harry) are not among A’s component haecceities.

The consensus seems to be that views like set-theoretical idealism were rightly discarded in history’s trashbin of misguided ideas. But it has great explanatory power. Below, I explain how set-theoretical idealism handles all six of the questions we wanted to answer within the context of a conservative metaphysics of sets.

(i) Why are there no two sets with the same members?

Sets are collective ideas of haecceities, and each haecceity has a unique intentional object. A collective idea I₁ is identical to a collective idea I₂ just in case every component of one is a component of another. So if some collective ideas components are all haecceities, then no idea is identical with it unless it has all and only the same component haecceities. Haecceities are identical just in case their intentional objects are identical. Since a sets members are the intentional objects of its component haecceities, we can now explain why sets are identical when they have the same members. Set A has all the same members as set B just in case A and B consist of the same component haecceities. But if A is a collective idea with the very same component haecceities of the collective idea which is set B, then, by the identity conditions of collective ideas, set A = set B. Therefore, the truth of the Axiom of Extensionality falls out of the identity conditions for collective ideas and the tight connection between each thing and its haecceity.

(ii) How do sets “contain” both spatio-temporal and non-spatio-temporal entities in the same sense?

The spatio-temporal metaphor of containment is usually either left unexplained or analyzed in terms one thing’s being among other things in a plurality. The first is unsatisfying, since sets presumably contain spatio-temporal and non-spatio-temporal entities in the same sense. But how are we to understand the spatio-temporal metaphor in a way that captures the flexibility of the kinds of things sets contain? The Low Brow view of containment leads to difficulties with distinguishing sets from their members. set-theoretical idealism analyzes containment in terms of intention (or representation). More specifically, a set contains an entity when that entity’s haecceity is one of the
set’s component ideas. There are haecceities for all sorts of things: that something is spatio-temporal or non-spatio-temporal alone does not prevent someone from thinking of it. For each actual thing, whether the entity is a number, a set, a person, or whatever, as long as its haecceity exists, it is an admissible candidate to be a component of a collective idea alongside other haecceities. Sets contain spatio-temporal and non-spatio-temporal entities in the same way, via the intentional or representational content of their components.

(iii) What is the empty set, and why believe there is one?

The empty set is a collective idea with no component haecceities. It is the idea nothing. Since the empty set is the idea nothing (or the idea none), it has no component haecceities. And since it has no component haecceities, there is no one thing, or many things for that matter, that the idea is of or represents. Consequently, the set which is the empty set contains no individuals. Thus, the empty set is something, it is the idea nothing. But it contains no individuals—it is empty—because its content represents the absence of any individuals. Finally, set-theoretical idealism gives us good reason to believe there is such a thing as the empty set. Experience shows that we have the idea nothing. And we know the idea exists through introspection or perhaps even through remembering a time when we thought it.

About the empty set, Lewis (1991: 13) writes:

Must we then accept the null set as a most extraordinary individual, a little speck of sheer nothingness, a sort of black hole in the fabric of Reality itself? Not really. We needn’t be ontologically serious about the null set. It is useful to have a name that is guaranteed to denote some individual, but it needn’t be a special individual with a whiff of nothingness about it. Ordinary individuals suffice.

Lewis is at a loss to take the empty set at face value. He takes the singleton operation as primitive and has nothing left to identify with the empty set and also grant that it exists as a set and that it contains nothing. The great advantage of the present proposal is that the empty set is not a most extraordinary individual, a speck of sheer nothingness, or a sort of black hole in reality. We can be ontologically serious about it and take its standard definition, as a set which contains nothing, at face
value. It is as real as any other idea. Its “whiff of nothingness” resides in its representational content, not in its very nature.

(iv) Why are singleton sets distinct from their lone members?

A singleton set is a single haecceity. The singleton set \( \{ \text{Tom} \} \) is the haecceity \( = \text{Tom} \). \( \{ \text{Tom} \} \) contains Tom because the haecceity which it is, is Tom’s. Furthermore, \( \{ \text{Tom} \} \) is distinct from Tom because the former is an idea of Tom and Tom is not an idea of Tom. Or consider the singleton set of the empty set, \( \{ \emptyset \} \). Even \( \emptyset \) has its own haecceity, \( = \emptyset \) the idea of being it. In this case, too, \( \emptyset \)’s haecceity is distinct from the idea which is \( \emptyset \). The idea which is \( \emptyset \) is the idea \textit{nothing} whereas \( \emptyset \)’s haecceity is the idea of \textit{that} (where \textit{that} is itself the idea \textit{nothing}, \( \emptyset \)).

Set-theoretical idealism allows us to take set-theoretical language about singletons at face value without rejecting that language or reinterpreting it. Moreover, it captures the language of singletons without appealing to an abstract container whose metaphysical nature is unclear at best. This gives set-theoretical idealism an explanatory advantage over its competitors.

It is difficult to overestimate how much an advantage set-theoretical idealism has in this regard. Lewis rightly observes that “the notion of a singleton was never properly explained,” and wonders “how it is possible for us to understand the primitive notion of singleton, if indeed we really do.”\textsuperscript{45} And he rightly claims that this ignorance about the nature of singletons infects our understanding of sets in general:

We were told nothing about the nature of the singletons, and nothing about the nature of their relation to their elements. That might not be quite so bad if the singletons were a very special case. At least we’d know about the rest of the classes. But since all classes are fusions of singletons, and nothing over and above the singletons they’re made of, our utter ignorance about the nature of the singletons amounts to utter ignorance about the nature of classes generally. .... What do we know about singletons when we know only that they are atoms, and wholly distinct from the familiar individuals? What do we know about other classes, when we know only that they are composed of these atoms about which we know next to nothing?\textsuperscript{46}

\textsuperscript{45}Lewis (1991, vii).

\textsuperscript{46}Lewis (1991, 31).
We needn’t accept Lewis’s view that sets are mereological fusions of singleton sets to see that ignorance about singletons is a symptom of a more pervasive ignorance about sets generally. Here is the main question: what is the relation between a set and its members? Insofar as we cannot answer this question for singletons, we cannot answer this question for sets generally. But rather than reject singleton sets as others have done in various ways, Lewis asks us to

...subdue our scepticism, and have faith in the teachings of set-theoretical mathematics.

Let us accept the orthodox iterative conception of set, including the part of it that escapes elementary formulation.\(^47\)

He then proceeds to build an entire theory on a primitive notion of singleton which he claims not to understand. Set-theoretical idealism has a more sure foundation.

(v) If sets depend on their members metaphysically, how should we understand that dependence?

There is no idea of being that unless there is a “that” to which the mind can direct itself. A haecceity is an haecceity of something actual, and so without the something, it’s haecceity does not exist. Haecceities are parasitic on the existence of their actual intentional objects. Thus, the metaphysical dependence of a set on its members arises from the dependence of each haecceity on its intentional object. But there are other features of this dependence which ground the well-foundedness of sets.

First, haecceities are also essentially directed towards something besides themselves. (Try to think something like that idea so that what is thought is itself the object of the thought. It cannot be done.) A haecceity cannot be its own intentional object. As a result, no set is self-membered.

Second, in order for there to be a membership loop in the hierarchy of sets, some set would have to be a member of one of its members, or one of its members’ members, or—and so on. This is not possible on the idealist view. Let’s consider a simple membership loop to diagnose why set-theoretica idealism precludes the existence of membership loops. Suppose set S has two members, A and B. S will then be the collective idea \(=A\) and \(=B\). Now suppose A is also a set, \(\{S\}\). A will then be the haecceity\(=S\). Then, what are the intentional objects of \(=A\) and \(=S\)? Well, the intentional object of \(=A\) should presumably be A, the set \(\{S\}\). But the set \(\{S\}\) is the haecceity \(=S\). What is the intentional object of \(=S\)? Why, it is S, the collective idea one of components is \(=A\). In order to have a

\(^{47}\)Lewis (1991).
membership loop, we will have to put on hold, as it were, what the intentional objects are of some haecceities until we have formed the loop. But haecceities do not put their intentional objects on hold in this way. Haecceities have their objects from the get-go and they are well-defined.

Third, in order for there to be an infinitely descending membership chain, there would have to be a collective idea of haecceities one of whose components was an haecceity of a collective idea of haecceities, one of whose components was an haecceity of a collective idea of haecceities, and so on, without end. The problem here is that each haecceity in the infinitely descending chain wouldn’t seem to have a determinate intentional object—what each haecceity would be an idea of would depend on what some prior haecceity is an idea of, and since there is no end, no well-defined intentional object of an haecceity at the ground-level, no haecceity in the chain would have a well-defined haecceity. The content of any such haecceity would depend on the content of another. And another. And another. Haecceities don’t work that way. So we have good reason to believe that if set-theoretic idealism is true, sets depend on their members in a way that would secure well-foundedness.

(vi) What kind of epistemic access do we have to sets?

If sets are ideas and ideas are the raw materials of thought, we should have epistemic access to a great many sets. The sets we fully grasp are those whose number of haecceities does not exceed our cognitive limitations. We should be able to grasp the empty set, various singleton sets, and so on, until our cognitive powers are too limited to grasp all the components of a collective idea or all the iterations of the ‘set of’ operation. The benefit of placing set-theoretical space within mental space is that it automatically explains our access to a vast swath of the lower reachers of the cumulative hierarchy.

The cost of placing set-theoretical space within mental space is the challenge of explaining (i) why there are enough ideas to identify with sets, given that ideas within the human mental community are much too few to account for the higher reaches of the cumulative hierarchy, and (ii) how we know of sets in those higher reaches of the hierarchy even if we cannot fully grasp them. For the simplest and most straightforward solution to (i) and (ii) is to posit an infinite mind which houses the set-theoretical universe and shares those ideas. Rather than develop that view fully here, I’ll direct the reader to Sections 14 and 15 in Chapter 4. On that view, ideas are mind-dependent objects which channel the efficacy of an infinite mind which holds them in being. That mind’s boundlessness
ensures that there are enough ideas. And the causal efficacy of ideas is a natural extension of the view that the raw materials of thought ultimately belong to an omnipotent being with an infinite mind. Those with much more naturalistic tendencies may think such a view is unlikely to be true. But the view that there is a causal connection between the infinite ideas of an infinite mind and our minds would, if true, partly dispel the mystery of our epistemic access to so-called abstract objects like sets. For if there is not a causal connection between us and the set-theoretical universal, what kind of connection could there be which would also explain how we know as much as we seem to know?
CHAPTER 7
Numerical Idealism

7.1 Introduction

Cardinal numbers say how many things there are in a collection. If you have cats, the cardinal number of your cat collection says how many cats you have. What are these cardinal numbers, metaphorically speaking? Given their metaphysical nature, whatever it is, how do we know that they exist, understand them, and reliably apply them to the collections which they number?

These questions are more tractable if we restrict ourselves to the numbers with which we are most familiar, the numbers 0, 1 and 2. Since an account of numbers should capture both what we seem to know about them and how we could know it, our familiarity with 0, 1, and 2 may guide us towards a correct account of the natural numbers generally. So we may reframe our original questions concerning what the cardinal numbers are and how we know about them as questions about what the numbers 0, 1, and 2 are and how we know about them.

We cannot exploit our familiarity with 0, 1, and 2 to explain how we know about them without first pinpointing some known familiar facts about them. One kind of familiar fact is, as Dummett says in a slightly different context, “so simple that it needs a sophisticated intellect to overlook.” The number 2 numbers pairs of things, whether we use ‘1’ or ‘2’ or ‘3’ as its name or whether we use ‘2’ in some contexts to number one thing or three things. A pair of things consists of an \( x \) and a \( y \) (which are distinct) and no other things. If \( x \) and \( y \) are distinct, the number 2 numbers them. But

\[ 1 \]

Hereafter, I will drop the ‘cardinal’ in ‘cardinal number’.

\[ 2 \] Dummett (1991, 53). Dummett’s remark here appears in an argument against Dedekind’s structuralist account of natural numbers and in favor of Frege’s strategy of making the application of the natural numbers as cardinals central to their definitions. This is not the point I wish to endorse. The point is simply that cardinal numbers do number collections—what explains their ability to do this?
if $x$ is identical to $y$, the number 1 does. And 0 is the number of things which are not identical to anything, namely, nothing.

I have two kinds of questions about this connection between number and identity. First, why are the numbers 0, 1, and 2 tied to the notions of identity and diversity in the ways they are. What about the number 2, for example, explains why it alone numbers pairs of things? And why do 1 and 0 alone number single-membered and empty collections, respectively? Second, how are we reliable judges about the cardinalities of smaller collections in the way we seem to be? Is there a feature of the number 2 we grasp which also explains why it alone numbers pairs of things?

In this paper, I rehabilitate a traditional but unpopular account of the nature of cardinal number within a traditional but unpopular metaphysical framework. The traditional account is the units view which says that each natural number is a collection of indistinguishable units. The metaphysical framework places mathematical reality within mental space. Partly due to Frege’s withering criticisms in Sections 28–44 of the Grundlagen, criticisms which Michael Dummett has called “conclusive,” the units view has very few contemporary defenders. And Frege’s arguments against mixing math and psychology are also partly responsible for the dearth of idealist views of mathematical objects. But the numerical idealism I offer survives Frege’s criticisms and answers the above questions about the numbers 0, 1, and 2. For these reasons and others, I argue that an idealist units view deserves further consideration.

## 7.2 Units

Ancient mathematicians often used tallies to represent number. One tally stood for the number 1, two tallies for the number 2, and so on, until the tallies gave way to a more convenient symbol. The ancient Egyptians and Babylonians used tallies up through the numbers 3 and 9, respectively, and the Pythagoreans used pebbles or dots in the sand to similar effect. Eventually, numbers were elevated to denizens of an ideal world of abstract objects, and the tallies which represented numbers in concreto were thought to reveal an essential aspect of their internal structure. Whereas numerals might consist of collections of indistinguishable tallies, each number was a collection of indistinguishable units,

---


4 Kline (1990, 5, 15, 30).
a collection of heavenly tallies. Two indistinguishable tallies signified the number 2, for example, which, in turn was thought to consist of two indistinguishable units.

There are two main reasons to hold the units view. First, the account is intuitive, at least pre-theoretically. For suppose we identify the natural numbers with sums of 1s and then identify the 1s with indistinguishable units. Then, that 1 = 1, 2 = 1 + 1, 3 = 1 + 1 + 1, 4 = 1 + 1 + 1 + 1, and so on, implies that each natural number $n$ consists of (a set of) $n$ units.\(^5\) More recent attempts to construct the natural numbers out of sets define 1 as the successor of 0, 2 as the successor of 1, and so on. At the very least, these set-theoretic characterizations of the natural numbers are now thought to capture essential structural features of the natural numbers, even if they fail as true definitions.\(^6\)

What are these structural features which sets seem to capture? It isn’t too far fetched to think that the set-theoretic characterization of 1 as 0’s successor captures 1’s being a single unit; that the set-theoretic characterization of 2 as 1’s successor captures 2’s consisting of the unit which is 1 plus an additional unit; that the set-theoretic characterization of 3 as 2’s successor captures 3’s consisting of the two units which comprise 2 plus an additional unit; and that, generally, a successor to any number is the totality of units of its predecessor plus another unit.

Second, the units view is supposed to explain why equinumerous collections (collections whose members can be paired off without remainder) have the same number of individuals, what Dummett (1991: 145) calls the “whole point” of the theory. Suppose I can pair each left-foot shoe with a right-foot shoe and put them in the closet together without having any remaining shoes left. Then, the number of left-foot shoes is the same as the number of right-foot shoes. The units view is supposed to explain why, but this explanatory virtue has less to do with the notion of a unit and more to do with a crucial component of the units view called abstraction. How does abstraction work, and how does it explain why equinumerous collections have the same number of individuals?

Unsurprisingly, there are different views of abstraction. So before we answer these questions, it will be beneficial to survey three historically important views of abstraction. Husserl, Cantor, and Dedekind each subscribed to a different view of abstraction. For each, abstraction is a mental process which selectively ignores the features of some object or objects. Husserlian abstraction takes some

\(^5\)Though, as Frege (1884, Sect. 6), these definitions rely on the associativity of addition. We might want an account of number out of which the associative law falls, not one on top of which we must assume the associative law.

\(^6\)Benacerraf (1965).
collection or “totality” of things and provides the “ones” or units which constitute the number of that totality. Husserl explains:

To each such [concrete] totality there “belongs a certain number.” It is easy to characterize the abstraction which must be exercised upon a concretely given multiplicity in order to attain the number concept under which it falls. One considers each of the particular objects merely insofar as it is a “something,” or “one,” simultaneously retaining the collective combination; and, in this manner, there is obtained the corresponding general form of multiplicity, one and one and ... and one, with which a determinate number name is associated. In this process there is a total abstraction from the specific characters of the particular objects. But this neither means nor implies that the concrete objects have to disappear from our consciousness. To “abstract” from something merely means to pay no special attention to it. Thus, also in our case at hand, no special interest is directed upon the peculiarities of content in the separate individuals, while those peculiarities, nonetheless, do constitute the pre-condition of the acts of reflexion which yield the “units” of the respective number, and are the ground of the distinctness of those units.7

For Husserl, abstraction is a mental process which begins with the consciousness of some collection of things and arrives at a collection of units. We strip down each thing to a kind of mental bare particular by ignoring all of its properties except for its distinctness from other objects in the collection. The result is a combination of units which constitutes the number of the collection with which we originally began. So if we begin with a collection of four goats, say, one may selectively ignore everything except for their mere diversity, the fact that each is different from the others. The result is a combination of four indistinguishable units, the number 4, which is the number of goats with which we began.

Cantorian abstraction involves an additional step. Like Husserl, we begin with some collection of objects. Cantor calls each such collection an “aggregate” (Menge), which is “any collection into a whole (Zusammenfassung zu einem Ganzen) M of definite and separate objects m of our intuition or our thought.”8 Then abstraction takes us from each aggregate M to the number of its elements:

8Cantor (1955, 85).
Every aggregate $M$ has a definite “power,” which we will also call its “cardinal number.” We will call by the name “power” or “cardinal number” of $M$ the general concept which, by means of our active faculty of thought, arises from the aggregate $M$ when we make abstraction of the nature of its various elements $m$ and of the order in which they are given. We denote the result of this double act of abstraction, the cardinal number or power of $M$, by $\overline{M}$. Since every single element $m$, if we abstract from its nature, becomes a “unit,” the cardinal number $\overline{M}$ is a definite aggregate composed of units as an intellectual image or projection of the given aggregate $M$.\(^9\)

Later on, Cantor defines an aggregate $M$’s “ordinal type” $\overline{M}$ as the “general concept which results from $M$ if we only abstract from the nature of the elements $m$, and retain the order of precedence among them.”\(^10\) Thus, we may think of an aggregate $M$’s cardinal number, or number, as the result of a two step process. In the first step, we abstract from the nature of its elements but not from their order. This abstraction results in the ordinal type $\overline{M}$. In the second step, we then also abstract from the order retained in $\overline{M}$. The final result is the unordered aggregate of pure units, $\overline{M}$, the number of elements in $M$.

Whatever Cantor’s view of abstraction really was, we needn’t think of Cantorian abstraction as a temporal process whose first step occurs at one time and whose second step occurs at some later time. We might instead construe the two steps as simultaneous, if we insist on abstraction as being temporal at all. For if each number is a set of indistinguishable units, then there is no order among the units to distinguish them. Abstraction from a collection of objects to a collection of pure units as the cardinal number of that collection rids the original collection of any order implicit among its elements. So this construal of Cantorian abstraction fills out what must of been, at the very least, an implicit feature of Husserlian abstraction.

Let us pause here to note how Husserlian or Cantorian abstraction is supposed to explain why equinumerous collections have the same number of individuals.\(^11\) Since abstraction merely replaces each element of a set with a bare unit, we may represent abstraction as a function from any set $A$ to a set $C$ which is equinumerous with $A$ and whose elements are distinct units. Call it the $a$-function.

\(^9\)Cantor (1955, 86).
\(^10\)Cantor (1955, 111-112).
\(^11\)The explanation here is similar to the one in Dummett (1991: 154).
Does the a-function give the very same set of bare units C as the value of any equinumerous sets A and B? What we have to rule out is the possibility that the a-function gives equinumerous though distinct sets of bare units as the values of equinumerous sets A and B. For example, suppose we begin with a set of goats and a set of books and that we can pair the goats and books without remainder. If we want to explain why these two sets have the same number of individuals given that the number of each is a collection of bare units, then abstraction better not give us one set of bare units for the set of goats and another set of bare units for the set of books. The defender of the units view must provide us some reason for thinking that the result of abstracting from either of any pairs of equinumerous sets will be the very same set of bare units. This is not a trivial task. On the units view, there are infinitely many indistinguishable units. Do we attain the very same set of bare units no matter which set of n objects we abstract from? This is the equinumerosity problem. A units view should explain why equinumerous collections have the same number of members.

Dedekind’s characterization of the natural numbers raises an important question about the connection between mathematics and psychology. He uses abstraction to define all the natural numbers in one fell swoop, and, for that reason, seems more dissimilar to Cantorian and Husserlian abstraction than either of the latter two are to each other. Dedekind defines a simply infinite system in terms of one of its elements e and a one-to-one function φ whose domain and range are subsets of a set M:

(i) e is not the value of φ(α) for any element α of M,

(ii) the value of φ(α) for any element α of M is some element β of M, and

(iii) M is the least or smallest set satisfying conditions (i) and (ii).

In other words, M is the least set which has a base element e and is closed under φ, which will be something like the successor function. Given a set M which meets conditions (i)-(iii), Dedekind defines all the natural numbers at once as the result of abstracting from M:

73. Definition. If, in considering a simply infinite system [M] ordered by the mapping φ, we completely disregard the particular nature of the elements, retaining only their distinguishability and considering only those relationships in which they are placed to one another by the ordering map φ, then these elements are called natural numbers.
or simply numbers, and the base element \([e]\) is called the base element of the number series \([M]\). In consideration of this freeing of the elements from every other content (abstraction) one can with justice call the numbers a free creation of the human intellect [des menschlichen Geistes].

Dedekindian abstraction identifies numbers with positions in an abstract structure common to any simply infinite system which meets conditions (i)-(iii). Though he does not call numbers “units” here, it would not be inaccurate to do so. As William Tait (1997: 76) observes, abstracting from a simply infinite system to the system of numbers seems to involve nothing more than the first step of Cantorian abstraction writ large. One begins with an ordering of objects and then abstracts from the features of those objects while preserving not only their diversity but also their order. The numbers are, in a sense, ordered units.

In the passage above, Dedekind calls the numbers a “free creation of the human intellect.” Does Dedekind mean something else by this or does he think the numbers really are creations of the human mind? We can ask similar questions or Cantor and Husserl. For Cantor calls the number of a collection \(M\) an “intellectual image or projection of the given aggregate \(M\),” and Husserlian abstraction is clearly a psychological process. Some of Frege’s main criticisms against the units view maintain that psychology has no place in the definition of number. It is controversial to what extent Cantor, Dedekind, and Husserl would have disagreed with Frege. Whatever they actually thought, Frege’s scathing criticisms of the units view have generally been regarded as sound and contemporary defenders are few and far between. Even the defenders who remain still take Frege’s criticisms against psychologism to heart. This was a mistake: the most explanatorily powerful version of the units view is wholly psychological. I will present that account later on. In the next section, I rehearse Frege’s most influential arguments against the units view so that, in the subsequent section, I can explain how the more recent, non-psychologistic versions respond to those arguments.

\(^{12}\)Dedekind (1963, 68).

\(^{13}\)See n. 3.
7.3 Frege’s Criticisms

Frege launches a sustained attack on the units view in § 28-44 of his *Grundlagen*. There, Frege criticizes both the process of abstraction and the notion of a unit.¹⁴ Let’s consider Frege’s criticism of abstraction first. In § 34, he focuses on the views of his contemporaries Johannes Thomae and Rudolph Lipschitz:

For suppose that we do, as Thomae demands, “abstract from the peculiarities of the individual members of a set of items,” or “disregard, in considering separate things, those characteristics which serve to distinguish them.” In that event we are not left, as Lipschitz maintains, with “the concept of the Number of the things considered”; what we get is rather a general concept under which the things in question fall. The things themselves do not in the process lose any of their special characteristics. If, for example, in considering a white cat and a black cat, I disregard the properties which serve to distinguish them them, then I get presumably the concept “cat.” Even if I proceed to bring them both under this concept and call them, I suppose, units, the white one still remains white just the same, and the black black. I may not think about their colours, or I may propose to make no inference from their difference in this respect, but for all that the cats do not become colourless and they remain different precisely as before. The concept “cat,” no doubt, which we have arrived at by abstraction, no longer includes the special characteristics of either, but of it, for just that reason, there is only one. (S. 34)

Frege here argues that abstraction isn’t up to the task required by its defenders. The units view says that we arrive at a collection’s number, a collection of distinct though indistinguishable units, through a process of abstraction which ignores the distinguishing characteristics of the original collection. But once we abstract away from the distinguishing characteristics of the original collection’s members, we are left with a single general concept under which each of those members falls. For Frege, abstraction from the concept of a white cat and a black cat goes off the rails before we get to

¹⁴Kit Fine (1998: 604 ff.) considers two objections against Cantorian abstraction, one against the notion of abstraction and one against the notion of a unit. But Fine admits that he has “not attempted to determine the exact relationship between my own objections and [Frege’s].” This is admirable, I think, since Frege seems to mischaracterize the views he criticizes. But the arguments Fine considers dovetail nicely with my understanding of Frege’s criticisms, even if Frege’s criticisms mischaracterize the views he criticizes.
the concept of two pure units. As we abstract from the distinguishing characteristics of each cat, we
arrive at the general concept of a cat. And we arrive there whether we begin with two cats, four cats,
or six. If we kept going, we’d arrive at a unit, but only one.\textsuperscript{15} The lesson Frege wants to impart is
this: since abstraction provides the same result for variously sized collections, we cannot rely on it to
provide us with a general concept of number.\textsuperscript{16} This is the \textit{abstraction problem}. Abstraction from a
set of individuals either provides an equinumerous set of impure units or a singleton set of a pure
unit.

In sections 36-38, Frege turns his sights on the notion of a unit. His first main argument against
units begins with a short quotation from W.S. Jevons:

Whenever I use the symbol 5 I really mean

\[ 1 + 1 + 1 + 1 + 1, \]

and it is perfectly understood that each of these units is distinct from each other. If
requisite, I might mark them thus

\[ 1' + 1'' + 1''' + 1''''. \]

It is “perfectly understood” that the units are distinct from one another because a unit added to itself
is 1 unit and not 2 units. So if we signify any of the units in ‘1 + 1 + 1 + 1 + 1’ more than once,
the formula itself will signify a number less than 5. Frege agrees with Jevons that we should mark
units differently, given the units view along with the view that the units are distinct. For if the units
are distinct but we mark them similarly, we wouldn’t know for sure whether, for example, we’ve
signified a single unit twice or two units once apiece. To avoid difficulties like this, we would have

\textsuperscript{15}The first objection Fine (1996: 604) considers “amounts to a proof that there can only be two numbers under the
Cantorian account: 0 and 1.” Fine defends Cantorian abstraction against the criticism. But Deiser (2010, 135) endorses
the criticism when he says that “the indistinguishable units get lost by the axiom of extensionality” in ZFC.

\textsuperscript{16}Michael Hallett (1984: 130) also pinpoints the above paragraph from S. 34 as containing one of Frege’s two main
criticisms of the units view. Hallett’s only comment on the paragraph is this: “Thus intellectual ‘transformation’ of one
object \( (m) \) into another \( (E) \) is impossible” where \( m \) is the element in the original set to be abstracted from and \( E \) (for \textit{Eins})
is the one or unit we abstract to. Hallett misses the main point here. Frege’s main point isn’t that we cannot abstract from
a thing to a unit or that we cannot metaphysically transform a cat into a unit by thinking. Hallett’s quotation leaves out
the last sentence of the above quotation, the conclusion to the section which is titled \textit{Are units identical (gleich) with one
another?}. That sentence says that the result of abstracting from two cats leaves us with the very same general concept.
to adopt a general rule for when occurrences of ‘1’ signify different units. Frege considers the rule that each occurrence of ‘1’ in a statement signifies a different unit. But such a general rule has the unfortunate consequence of falsifying ‘1 = 1’. For if each unit is distinct, they are not the same. Frege concludes that in the absence of a rule for when occurrences of ‘1’ signify different units, the units view requires something like Jevons’ technique of distinguishing units.

Yet if we mark the units differently to represent their distinctness, we face the following challenge from Frege. If the units which compose the number 5 are distinct, why should we prefer Jevon’s formula, ‘1′ + 1″ + 1‴ + 1⁴ + 1⁵’, to an expression such as ‘a + b + c + d + e’? Frege’s challenge hinges on the claim that if we “assign different symbols to different things, it is hard to see why we still retain in our symbols a common element…” Frege’s challenge threatens the very indistinguishability which supposedly allowed them to represent number in the first place. We’ll call this the plurality problem.

The plurality problem may also be put in the form of a dilemma. If the number 3 consists of three 1s, then each 1 is distinct from each of the others. Distinguishing 1s in this way leads to multiple 1s. Since there seems to be no prohibition on grouping the different units from larger numbers to form smaller combinations of units, all the different groups of units lead to different 2s and 3s and 4s and so on. This proliferation of number 1s, 2s, and so on, is “utterly incompatible with the existence of arithmetic.”

But if there is only one number 1, then how could there be numbers greater than 1, which are supposed to consist of multiple 1s? Frege concludes:

The symbols

\[ 1′ + 1″ + 1‴ \]

tell the tale of our embarrassment. We must have identity—hence the 1; but we must have difference—hence the strokes; only unfortunately, the latter do the work of the former.

---

17 Frege (1884, 47).
18 Sutherland (2008, 6) calls this the “pure plurality problem.”
19 Frege (1884, 49, 57).
20 Frege (1884, 57). Fine (1998: 606) considers a very similar objection. There he wonders why some units should be especially suited to compose a number or to be the products of abstraction and not others.
21 Frege (1884, 47-48).
William Tait (2007: 99) claims that Jevons and Frege here “collaborate in a confusion.” The confusion, according to Tait, is that “if cardinals are sets, then their addition is just their union.” If cardinal numbers are sets, the sum of cardinal numbers \( m \) and \( n \) is not their union but their so-called “disjoint union.” For if \( m \) and \( n \) have at least one member in common, then the cardinality of their union will be less than the sum of their cardinalities. The cardinality of their sum is not the cardinality of their union, then, but the cardinality of their disjoint union, which we can construct in the following way. Take each member of \( m \) and form an ordered pair with it and 0. The resulting set of ordered pairs will have the same cardinality as \( m \). Then, take each member of \( n \) and form an ordered pair of it and 1. The resulting set of ordered pairs will have the same cardinality as \( n \). But now the two resulting sets are disjoint. So the cardinality of their union will equal the sum of \( m \)’s and \( n \)’s cardinalities. Tait then writes:

Frege (§ 38) somewhat misses the point here: He thinks that Jevon’s problem arises from confusing ‘unit’ with ‘one’ and of treating numbers as ‘agglomerations’. He is right that one cannot take the units in 5 to be the unit in 1, since there are five of the former and only one of the letter. But Jevon’s equation \( 5 = 1 + 1 + 1 + 1 + 1 \) leads to this identification, not because he thought that numbers are sets (after all, the von Neuman cardinals are sets), but because he confused the addition of cardinals with the union of sets.

Tait recognizes the force of Frege’s point that the units in 5 cannot be the unit which is 1, but Tait believes the ultimate problem with Jevons’ view is that he confuses addition and set-theoretic union. Yet if we have our eye on the units view, as Frege does, then the ultimate problem Frege has with the units view is the following. On the one hand, units must be distinct if numbers greater than 1 consist of them. But, on the other hand, units must be identical if there is only one number 1. The unfortunate representation of a combination of units as a sum of 1s is not necessary for Jevons or Frege to make their respective points. Instead, they might have said that, according to the units view, 5 is 1 and a 1 and a 1 and a 1 and a 1 without using ‘+’. Then Jevons could go on to insist that since 5 consists of five different units, we should represent 5 as \( 1' \) and \( 1'' \) and \( 1''' \) and \( 1'''' \) and \( 1''''' \). And in response, Frege would still insist that the distinctness of units poses difficulties for the units.
view. Focusing on the problems with cardinal arithmetic here, as Tait does, distracts from Frege’s philosophically important criticism of the units view.

Frege’s criticisms surveyed so far focus on the twin pillars of the units view, the notion of a unit and the process of abstraction. Frege also criticized the units view for mixing mathematics with psychology. We can separate at least two kinds of criticisms against psychologism, one against each pillar of the units view. The first criticism against the intrusion of psychology into mathematics concerns units themselves as psychological objects. The criticism is that psychologism about mathematical objects threatens the objectivity of mathematics. Frege uses ‘idea’ to mean private mental objects. So if numbers were ideas, we should have to speak of many twos. But in spite of the many millions of twos, there would not be enough ideas with which to identify all the numbers we ordinarily suppose there are.

The second criticism concerns abstraction as a psychological process. Frege’s most colorful presentation of this criticism appears in a draft of a review of Cantor, but this criticism does not appear in the published version. Frege writes:

So let us get a number of men together and ask them to exert themselves to the utmost in abstracting from the nature of the pencil and the order in which its elements are given. After we have allowed them sufficient time for this difficult task, we ask the first ‘What general concept have you arrived at?’ Non-mathematician that he is, he answers ‘Pure Being’. The second thinks rather ‘Pure nothingness’, the third—I suspect a pupil of Cantor’s—‘The cardinal number one’. A fourth is perhaps left with the woeful feeling that everything has evaporated, a fifth—surely a pupil of Cantor’s—hears an inner voice whispering that graphite and wood, the constituents of the pencil, are ‘constitutive elements’, and so he arrives at the general concept called the cardinal number two. Now why shouldn’t one man come out with one answer and another with another? Whether in fact Cantor’s definitions have the sharpness and precision their author boasts is accordingly doubtful to me.\(^\text{23}\)

\(^{22}\)See Frege (1884, 33-39).

\(^{23}\)Frege (1979, 70).
Frege imagines a scenario in which various people are asked to conduct Cantorian abstraction on a single pencil. Then each person abstracts to a different concept. In fact, one person abstracts from a different aggregate to begin with, not the pencil but the graphite and wood. The problem with abstraction seems to be that which collection one begins with and which concept one ends with depends on the perspective of the one who conducts the abstraction. For Frege, Cantorian abstraction is too imprecise to play such an important role in an account of number.\[^{24}\] Given the view that mathematics concerns an autonomous, purely formal world of abstract objects, it is admittedly odd to think that a definition of number should include a (human) psychological process. To avoid these two problems, Fine attempts to strip any mental language from his account of Cantorian abstraction.

7.4 Finean Abstraction

Cantorian abstraction seems inherently psychologistic. Kit Fine (1998: 604) admits that once Cantor’s views are “stripped of their psychological trappings, nothing much seems to remain.” Yet Fine finds enough inspiration in Cantor to develop a non-psychologistic version which seems to withstand the criticisms of the previous section. For Fine, abstraction is not a psychological process and units are not psychological objects. In this section, I summarize Fine’s view. In the next section, I evaluate it.

Fine identifies units with the variable (or arbitrary) objects from his *Reasoning with Arbitrary Objects*.\[^{25}\] To help us understand what variable objects are, Fine compares and contrasts variable objects with variables. A variable is a linguistic sign which takes objects as values. Unlike variables, variable objects are non-linguistic objects. But like variables, they take other objects as values. Variable objects are “abstract objects which assume those values by way of of their intrinsic nature.”\[^{26}\] Fine does not explain what a variable object’s intrinsic nature is which allows it to assume or take other objects as values. Nor does he explain what it means for an object to assume other objects as values.


\[^{25}\] Fine (1998, 608–9) says that it is worthwhile to note that his theory of arbitrary objects was created with “completely different applications in mind.” But it isn’t clear why it is worthwhile to note this. Does this more recent application of variable objects require Fine to amend some aspect of its original formulation?

Variables and variable objects also differ in how they take their values. Variables take their values independently of one another—that any variable takes some value does not constrain the values other variables take. Fine’s variable objects often depend on one another for their values. To bring this difference in relief, it will be beneficial to adopt the following convention: lower-case, italicized letters such as $x$ and $y$ will be variables and the same, non-italicized letters $x$ and $y$ will refer to variable objects. For a case of dependence among variable objects, Fine claims that “there are [variable objects] $x$ and $y$ whose values are any real numbers $x_0$ and $y_0$ for which $x_0 = -y_0$.” We can think of the values of the [variable objects] $x$ and $y$ as “being simultaneously constrained by the requirement that the value of one should be the negative of the value of the others; or we may think of the value of one of the [variable objects], say $y$, as being given as the negative of the value of the other.”\(^{27}\) Though Fine deems one-way dependence as “somewhat more satisfactory,” he opts for the simplicity which two-way dependence provides.

Fine uses the notion of two-way dependence to provide an identity criterion for the sets of variable objects which he ultimately identifies with numbers. A set of variable objects is a two-way system just in case:

(i) any two variable objects in the set are co-dependent, and

(ii) no variable in the set co-depends upon any variable not in the set.\(^{28}\)

Fine then provides a criterion of identity for systems of variable objects. A system $S$ and a system $S'$ are similar just in case there is a one-one map from the variable objects of $S$ onto the variable objects which have the same values in $S'$. The criterion of identity says that the variable objects of similar systems are the same.\(^{29}\) This criterion rules out the scenario in which there are distinct though similar systems of variable objects. So, for example, if $S$ and $S'$ each consist of a single variable object whose values are the same, then $S$ and $S'$ are also the same, which implies that the variable object in $S$ is the same as the variable object in $S'$.

\(^{27}\)Fine (1998, 609).

\(^{28}\)Fine (1998, 609).

\(^{29}\)Fine (1998, 610).
With one more word about variable objects, we will be ready for Fine’s account of variable objects as units. He defines systems of variable objects with conditions on their values, and he states these conditions with uninterpreted variable signs. For example:

**LET** $x$ and $y$ be such that: $x$ and $y$ are reals and $y = -x$

defines a system of co-dependent variable objects $x$ and $y$ whose values are all real numbers $a$ and $b$ such that $b = -a$. Any system which meets the condition will be similar to any other. Since similar systems are identical, there will only be one such system. And there is only one such system even though there are two possible assignments of objects to variable signs ($x$ to $x$ and $y$ to $y$; and $x$ to $y$ and $y$ to $x$).

Fine identifies units with variable objects whose range is unrestricted. So a unit is a variable object whose values may be any object whatsoever. Each cardinal number is a set of units. And to specify which units a cardinal number is a set of, Fine appeals to two-way systems of variable objects. In such a system, units depend on each other for their values but do not depend on the values of any other units outside the system. Cardinal numbers are two-way systems of variable objects whose ranges are unrestricted.

Where $u_{mn}^i$ is a variable object with an unrestricted range, $m$ differentiates the unit from those of other systems, and $n$ differentiates the unit from those within its own system, Fine uses the following condition on the variable sign $u_{11}^1$ to define the number 1:

(1a) **LET** $u_{11}^1$ be such that: $u_{11}^1 = u_{11}^1$

(1b) **LET** 1 be such that: 1 = \{ $u_{11}^1$ \}

(1a) defines a unit and (1b) defines what the number 1 is, namely a singleton set whose single member is that unit. Fine then defines the number 2 with different units:

(2a) **LET** $u_{12}^2$ and $u_{21}^2$ be such that: $u_{12}^2 \neq u_{21}^2$

(2b) **LET** 2 be such that: 2 = \{ $u_{12}^2$, $u_{21}^2$ \}

---

Here, (2a) defines 2’s units and (2b) defines the number 2 as the set whose members are those units. Along these lines, Fine’s account will include similar definitions for the units $u_1^n, u_2^n, ..., u_n^n$ and, on the basis of those definitions, a definition of any natural number $n$.

Thus far, we’ve surveyed one half of Fine’s units view. The other half concerns abstraction itself. For Fine, the kind of abstraction which takes a set of objects and gives its number is a species of a more general abstraction operation. More generally, abstraction begins with the objects $a_1, a_2, ...$ and a relation $R$ represents the features of those objects we wish to preserve through abstraction. Those features we wish to preserve from $a_1, a_2, ...$ will be features of the objects $b_1, b_2, ...$, the objects which result from abstraction. If $R$ embodies those features we wish to preserve in abstraction, then abstracting from $a_1, a_2, ...$ will provide objects $b_1, b_2$ which themselves stand in the relation $R$.

Abstraction provides the number of a set of individuals when the only feature preserved is the relation of diversity, which holds of objects $b_1, b_2, ...$, just in case any two of them are distinct. When diversity is the only feature preserved in abstraction from a set of objects $a_1, a_2, ...$, the resulting objects are distinct variable objects $u_1, u_2, ...$ with unrestricted range—units, that is. The result of abstracting from a set of objects is an equinumerous set of distinct units. The units in the resulting set depend on each other for their values. So when abstraction preserves diversity alone, it takes a set of individuals and provides the systems of units which Fine has identified with numbers. Importantly, Fine denies that abstraction is a psychological process. So Finean abstraction does not necessarily help explain how we come to grasp numbers or justify our beliefs in their existence. Rather, “it will only be because we already know that there are objects with the desired feature that we can be assured that there exists an abstraction with respect to that feature.”

What Finean abstraction can explain is the relationship between equinumerous sets of individuals and the number which numbers them. This is one explanatory feature we sought in an account of number as (sets of) units, so let us now turn to evaluating Fine’s account.

---

7.5 An Evaluation of Finean Abstraction

How does Fine’s account fare as a species of the units view? Does it overcome the abstraction, plurality, and equinumerosity problems? Does it also capture what we want in an account of numbers?

Consider the equinumerosity problem, which is to explain why equinumerous collections of individuals have the same number of individuals. The problem with previous versions of the units view is that when we begin with distinct though equinumerous sets A and B, we have little reason to suppose the units abstracted from set A are the very same units abstracted from set B.

Fine’s identity criterion of similar systems implies that, given a class of equinumerous sets, abstraction provides the same set of variable objects. For the result of abstraction from a set A will be some system of variable objects and the result of abstraction from any set equinumerous with B will be some similar system of variable objects. What guarantees that the second system of variable objects is similar to the first is that their inputs are equinumerous. Since the systems of variable objects which result from abstraction are numbers, and since similar systems are identical, Fine’s identity criterion implies that equinumerous sets have the same number of individuals. Abstraction does not merely provide more equinumerous though distinct sets of variable objects or units.

What about the abstraction problem? Does Finean abstraction take any set of individuals and provide an equinumerous set of distinct units? By definition, it does. Finean abstraction takes any singleton set and provides the unique singleton set \{u_1\}. The latter is, of course, equinumerous with any singleton set. The number 1, \{u_1\}, is also Fine’s easiest case, since it consists of no other units from which it is supposed to be distinct. Consider the number 2, then. Finean abstraction takes any set of two individuals and preserves their diversity to provide a set of distinct units \{u_2^1, u_2^2\}. By preserving the diversity of the individuals in the original set, the resulting set of units will be equinumerous with the original set. Finean abstraction, therefore, takes any set of distinct individuals \{a_1, a_2\} and provides an equinumerous set of distinct units \{u_2^1, u_2^2\}. And, generally, Finean abstraction takes a set of distinct individuals \{a_1, a_2, ... , a_n\} and provides the unique and equinumerous set of distinct units \{u_n^1, u_n^2, ... , u_n^n\}.

Does Finean abstraction resolve the plurality problem? Are the units which comprise each number both indistinguishable and distinct? In *Reasoning about Arbitrary Objects*, Fine writes:
Arbitrary objects may be regarded as a special kind of Meinongian object. ... My arbitrary man, for example, is the Meinongian incomplete man. In general, an arbitrary object with certain values is the Meinongian object with the properties common to those values. Three important differences between the two theories should be noted though. First, my arbitrary objects are specified ‘from below’, in terms of the values they take; Meinongian objects are specified ‘from above’, in terms of the properties they have. This makes a big difference both to the formulation and development of the respective theories.34

Fine then mentions two other differences between his theory of arbitrary objects and Meinong’s theory of incomplete objects. But they are irrelevant to the present point. The present point is that since each unit is a variable object with an unrestricted range of values, and since the range of values determine what features a variable object has, the variable objects which are units should be indiscernible. But they are also distinct. Every unit within a number takes a different value from the values taken by every other unit within a number.35 And every unit within a number is distinct from every unit in every other number, since each number is an independent system of co-dependent variables. The unit in the number 1 is a unit which depends on no other unit for its value. The units in two depend for each other on their values, but do not depend on the value taken by the unit in 1 or the units in any number higher than 2. And so on. So we do not have multiple 1s, multiple 2s, and so on. Consequently, there is only one number 1, one number 2, etc.

Fine also intends his theory to be non-psychologistic. No psychologistic terms appear in his account of abstraction or in his definition of a unit. Fine offers his account of numerical abstraction as a set-theoretical function, not a psychological process. As a result, it evades Frege’s worry that both the initial collection and the resultant concept are relative to the person who conducts the abstraction. Furthermore, Fine’s units are variable objects, which he identifies with mind-independent abstract objects.36 It is clear that Fine would distinguish these abstract objects, which are publicly available, from private mental objects.

34Fine (1985, 44).
35Though each unit can take the same value as a unit in some other number. For instance, Barack Obama could be the value of a unit in 27 and also of the value of a unit in 184. Thanks to Bob Adams for this point.
Finally, we can evaluate how well the account explains a fact about the most familiar numbers. Take the number 2, for example. How does the view explain that when \( x \) and \( y \) are distinct, the number 2 numbers their collection and no other number does? For Fine, the distinctness of each individual from all other individuals in a set is the only feature preserved in numerical abstraction from a set to a set of units. If \( x \) and \( y \) are identical, then there is no distinctness preserved in abstraction, and so the resulting set will not contain distinct units. Similarly, if we begin with \( x, y, \) and \( z \), which are each distinct from the others, then abstraction yields an extra distinct unit. So it is only when we begin with a set consisting of something and something else and no other things that Finean abstraction yields the number 2, an independent system of co-dependent variables with unrestricted range.

However, central aspects of Fine’s view call for explanation, even if we grant that variable objects exist and take other objects as values and so on. The most serious worries concern variable objects themselves. Fine (1998, 609) claims that variable objects “take” other objects as values, that they “assume those values by way of their intrinsic nature.” What is their intrinsic nature? And how does a variable object’s intrinsic nature determine which objects are its values?

As far as I can tell, Fine (1998) does not say. But the theory of variable objects is most well-developed in his *Reasoning with Arbitrary Objects*, and there he provides some clues about how we should understand variable objects and how their nature allows them to take on other objects as values. Yet we ought to be careful about importing too much from Fine’s book about variable objects into his theory of units as variable objects. He himself says that it is “worth noting that the theory [of variable objects] was originally developed with completely different applications in mind.”\(^{37}\) So I will treat some remarks from his book as a possible suggestion for how we should understand variable objects without attributing them to his official theory.

The suggestion derives from a distinction between generic predicates and conditions, on the one hand, and classical predicates and conditions, on the other.\(^{38}\) Generic predicates include ordinary predicates such as ‘being odd’ and ‘being mortal’, and generic conditions are obtainable from generic predicates through the classical operations of quantification and truth-functional composition. Classical predicates, according to Fine, include “special” predicates such as ‘being an individual number’ and ‘being in the range of’, and classical conditions are obtained with them through the

---


\(^{38}\) Fine (1985, 13).
classical operations. Fine then adopts the following thesis about variable objects, where a is the name of a variable object a, and i is a variable ranging over individuals in a’s range:

\[(G3) \text{ for any generic condition } \phi(x), \phi(a) \text{ is true iff } \forall i \phi(i) \text{ is true.}\]

As it stands, (G3) will not work. Variable objects are abstract objects, and nothing Fine says prohibits ‘being abstract’ from counting as a generic predicate alongside ‘being mortal’. Since every variable object satisfies such a predicate we may reason from left-to-right that every value of every variable object is abstract. Since all units are variable objects, then, the values of every unit would be restricted to abstract objects, contrary to the definition of a unit as a variable object with unrestricted range.

Adopting another distinction between predicates or another kind of predication would solve this problem. Fine could borrow Terrance Parsons’ distinction between nuclear predicates (e.g., ‘being tall’, ‘being mortal’) and extra-nuclear predicates (e.g., ‘exists’, ‘is mythical’, or ‘is fictional’).\(^{39}\) Some predicates which all variable objects satisfy, such as ‘being a variable object’ and ‘being abstract’, are plausibly extra-nuclear predicates. Then Fine could restrict (G3) not only to generic predicates, but to nuclear predicates as well. Or Fine could adopt Edward Zalta’s distinction between encoding and exemplification.\(^{40}\) A variable object exemplifies the properties of being a variable object and being abstract—it does not necessarily exemplify all the properties its values exemplify (e.g., being green or being mortal). But a variable object encodes those generic properties its values exemplify. Then Fine could fix (G3) accordingly. Either strategy resonates with Fine’s earlier claim that variable objects are a kind of Meinongian object.

With either augmentation, we would then have some explanation for how variable objects assume their values by their intrinsic nature. A variable object either encodes or has as nuclear properties all the generic properties of its values. Given the class of green things, there is a variable object which encodes or has as a nuclear property of being green. Given the class of humans, there is a variable object which encodes or has as nuclear properties each of the properties every human has. And, most importantly, given the class of self-identical objects, there is variable object which assumes as values everything which is identical to itself. Any such variable object, by dint of being identical to itself, will have an unrestricted range of values.

\(^{39}\)Parsons (1980).
\(^{40}\)Zalta (1988, 15-19).
Although we now have some clue about the intrinsic nature of variable objects in virtue of which they assume their values, we do not yet have an explanation of how they assume or take those values in the first place. Fine’s antipsychologism prohibits two kinds of accounts about how variable objects have their values. First, the values of variable objects cannot depend on any intention of ours. So no variable object is linguistic entity to which we assign a range of values. But this was never an option for Fine anyway. The entire theory of variable objects is built on the assumption that they are distinct from the linguistic objects we call “variables.” Nor can variable objects be regular old objects of which we say things such as “Let this pencil stand for any object whatsoever.” Nor can variable objects be any sort of mental construction whose values we stipulate. In any of these cases, the values of variable objects depend on our intentions, and Fine’s anti-psychologism prohibits this dependence. Second, variable objects cannot be ideas whose values include the objects which have the features that the variable objects represent through their content. Fine’s account rules out both psychological accounts of variable units and psychological accounts of how they assume their values. If not by intention or mental or linguistic representation, then in what sense do variable objects assume or take their values?

Since variable objects satisfy, in some sense, some of the same conditions as their values, perhaps variable objects assume their values through mind-independent similarity relations. But any green thing seems more similar to green things than any variable object is which assumes green things as values. The green thing has a color which we can see, as do many other green things; the variable object is an abstract object which we cannot see. The variable object may lack location, shape, and so on. So in what sense of similarity could we say that variable objects are similar to their values which we could not also say of any of one of those values in relation to all the others? I don’t know how else Fine might explain what it means for an object to assume other objects as values in a way that would prohibit ordinary objects or singleton sets of ordinary objects from doing the same. I confess that I find the notion of an abstract object “taking other objects as values” to be completely mysterious unless this is a more complicated way of talking about intention or mental representation.\footnote{Tennant (1983, 88–89) also suggests that variable objects, “if admitted, might be better assimilated to the domain of cognitive psychology. . . . One might regard arbitrary objects as incomplete mental representations—not so much objects of thought as objects within thought, by means of which we reason about the world of objects.”} But that is precisely what Fine denies.
Psychologistic units views are not subject to this criticism. Fine’s units are supposedly apt for numbering because they each have an unrestricted range of values. But we don’t know what it means for objects to have values. Psychologistic units views are apt for numbering because either we intend to use them to number or they have psychological content which allows them to represent each and every thing individually. Intention and mental representation are also mysterious, but we know through experience that we intend and represent. They are known aspects of reality, though they are otherwise mysterious. Psychologistic accounts of units appeal to a known though mysterious aspect of reality. Fine’s units view doesn’t appeal to a known though mysterious aspect of reality but instead merely posits a new mysterious aspect of reality.

The second concern relates to units as abstract objects, which are typically thought to lack any sort of causal power. If units lack any sort of causal power, what kind of connection could we have to them which would provide us some way of knowing about them? We needn’t assume that every object of knowledge is causally connected to us. But if someone says there are some objects with which we have no causal connection, we should wonder what kind of connection we could have to them which would account for our apparent knowledge. And if we evaluate theories by their explanatory power, it is certainly a knock against a theory of objects about which we apparently know quite a lot if that theory renders our apparent knowledge inexplicable. Fine claims his units are abstract objects. But he doesn’t explain how we might know anything about them. Some think Husserl identified units with mind-independent abstract objects, but at least he had an entire story to tell about how we access them—the psychological process of abstraction. Fine’s purely logico-mathematical account of abstraction does not purport to have a similar epistemological payoff.

Some units views evade this second concern, not merely by adopting a psychological account of abstraction, but by adopting a psychological account of units themselves. On these views, units are mind-dependent objects, not mind-independent abstract objects. Since units are placed within mental space, no mystery arises about how we have a connection to them. What exactly these mental objects are is disputable. But that there are such mental objects, broadly and neutrally construed, is beyond reasonable inquiry. If there were no such things, we wouldn’t even be able to raise the question about whether there are such things. The difficulty arises for Fine because he takes units out of mental space and places them somewhere else, perhaps somewhere like Plato’s heaven. Doing so means that he loses out on the epistemological payoff of a more traditional units view.
There is also a more specific concern about Fine’s treatment of the number 0. Fine identifies 0 with the empty set, which follows from his account of abstraction. But why should 0 be the empty set? Fine identifies numbers greater than zero with sets of units. Consider the number 1, for example, which is the singleton set of a specific unit, a particular variable object with unrestricted range. This unit’s unrestricted range is what Fine uses to justify the identification of the number 1 with the singleton set of that unit. No feature of the set over and above those features of its member plays an ineliminable role in explaining what the number 1 is or why the number 1 numbers the individuals in any singleton set. Similar remarks apply to the number 2. We can define the units in the number 2 as Fine does, as co-dependent variable objects with unrestricted range whose values depend on no other units. But we do not need the set of them to construct Fine’s theory. We can refer to the units plurally, and if we want we can do so with a singular term like ‘2’. No feature of the set of two units over and above the two units plays an ineliminable role in explaining what the number 2 is or why 2 numbers the individuals in any set whose members are an x and y which aren’t identical. The set-theoretical machinery in Fine’s definition of the natural numbers is little more than a kind of useful but eliminable formal scaffolding on which all the meaningful parts of his theory rest. The set-theoretical machinery is only formally convenient for numbers greater than 0 since the mathematical essence of each number consists in its units.

Yet the set-theoretical machinery is absolutely essential for 0 itself. Rid zero of the set which contains its units and nothing is left. But zero is something, not nothing. But what could that something be? It cannot be the unit whose range is absolutely restricted and empty, which would, I think, make as much sense in Fine’s units theory. The result of Finean abstraction is supposed to preserve the diversity among some individuals in a set to provide that set’s number, an equinumerous set of units. But if we begin with the empty set then we either arrive at nothing at all if we do without the set-theoretical formalism or we arrive at the empty set. Well, 0 is something and not nothing, so the only option for Fine is to identify 0 with the empty set. What is merely useful formalism for some numbers is, in the absence of any other plausible alternative, the number itself in 0’s case. Fine’s theory introduces an arbitrary bifurcation in the natural numbers. Unlike numbers greater than 0 in Fine’s account, there seems to be nothing intrinsic to 0 which justifies categorizing it as a number alongside the others.
So the main criticisms against Fine’s units view are three. It isn’t clear what variable objects are because it isn’t clear what their essential function of having a range of values is. It isn’t clear we have any connection to them, even if they exist. And it isn’t clear why zero is a set even though numbers greater than zero do not seem to be tied essentially to sets. In ways that will soon become clear, all three problems are consequences of formulating a non-psychologistic units theory. In the next section, I offer a psychologistic units view which overcomes these problems.

7.6 Numbers Are Ideas

Numbers are ideas. Which ideas are they? And what are ideas, in the first place? I’ll answer these questions in reverse order.

When I think of Obama, for instance, there is, first, the token mental act. Somehow or other, at some place and time, I point my mind’s eye towards Obama. The token mental act is not what I’m calling an idea. At the other end of my thought is the intentional object, the thing my thought is about. In this case, the object of my thought is Obama, the man. Though ideas are sometimes the objects of an idea, an idea is not generally its own intentional object. So an idea is neither the token mental act nor the idea’s intentional object. The mental act is directed towards Obama through the intentional or representational content of the mental act. This content is an idea of Obama. Let ideas be the intentional or representational contents embodied within mental acts. The claim that numbers are ideas should then be understood as the claim that numbers are a particular sort of intentional or representational content.

Ideas have two kinds of structure which are relevant for identifying numbers with ideas. The first is a kind of parthood structure. Complex ideas have other ideas as parts. Take the complete idea of Obama as he is in the actual world, for example. This idea has many parts, including the ideas of being human, being the 44th president of the U.S., being born in Hawaii, having two daughters, and being married to Michelle. As I claim in Chapter 4, ideational parthood is a kind of mental conjunction. The complete idea of Obama is a large conjunction of ideas. It is the idea of being this and that and so on. The ‘and’ here signifies the relevant notion of conjunction.

---

42These would all be time-indexed in some way, of course.
An idea $I_x$ is an abstraction from $I_y$ when every conjunct of idea $I_x$ is also a conjunct of idea $I_y$ but at least some conjunct of $I_y$ is not one of $I_x$’s conjuncts. For example, the idea of being human is itself a conjunct of the complete idea of Obama, but obviously not every conjunct of the complete idea of Obama is a conjunct of the more general idea of being human. Hence, the idea of being human is an abstraction of the complete idea of Obama. So is the idea of being identical to something. No matter something’s nature, it is identical to itself and hence identical to something. So the idea of any particular thing includes not only that it is identical to itself but also identical to something. I’ll call the idea of being identical to something the idea of being something. The idea of being something is a mental conjunct of the idea of any particular thing. Since the idea of being something is a conjunct of the idea of any particular thing but the idea of any particular thing includes more ideas besides, the idea of being something is an abstraction of an idea of any particular thing.

The second kind of relevant structure is collective. Collective ideas are the mental analogue of plural referring terms. Whereas plural referring terms refer to some individuals, say, Tom, Dick, and Harry, a collective idea is an idea of individuals as distinct individuals. A collective idea is an idea of this thing and that thing and that other thing, and so on. The idea \textit{Tom and Dick and Harry} is a collective idea of Tom, Dick and Harry.

Consider a collective idea of Tom and Harry. The idea of Tom includes that he is not identical to Harry; the idea of Harry includes that he is not identical to Tom. Let’s subject this collective idea to two consecutive steps of abstraction. First, there is an abstraction from this collective idea which subtracts from the idea of Tom everything except for the idea of being something and the idea of being distinct from Harry. This resulting abstraction also subtracts any ideas within the idea of Harry which concern Tom. So in this first step, the idea of Harry no longer includes the idea of being not identical to Tom. Instead, the idea of Harry includes the idea of being not identical to the indeterminate something which is not identical to Harry. At this first step, the abstraction represents something (which isn’t identical to Harry) and Harry.

The second abstraction subtracts from the idea of Harry everything except for the idea of being something and the idea of being distinct from the other something. This second abstraction represents a thing and another distinct thing. This final abstraction is the idea of something and something.

\footnote{I introduce them more fully in Chapter 5, Section 6.}
else. The idea of something and something else is a collective idea. Each component includes the idea of being something. And each component includes the idea of being distinct from the other something. Notice that the collective idea represents the distinctness of a thing and another, not the distinctness of the ideas themselves. One component does not include the idea of being distinct from the other idea, but the idea of being distinct from whatever the other component represents. This idea of something and something else is the number 2, the number of individuals the collective idea represents or is about. And the number 2 is an abstraction of any collective idea of a particular thing and another particular thing.

Here are the first few numbers in the number series:

0 = the idea nothing

1 = the idea something

2 = the idea something and something else

Hence, ‘the number 0’ is another name for the idea nothing; ‘the number 1’ is another name for the idea something; and so on.

Each number is a collective idea of units. And each unit is an idea with incomplete content. First, consider once more the unit which constitutes the number 1. This unit consists of the idea of something. Like the incomplete idea of being a man in general, which is an idea of no particular man, the idea of being something is an incomplete idea of a thing in general. Unlike the haecceities from Chapter 5 whose content determines an object, there is no particular thing which is the object of the idea. Like the incomplete idea of being a man in general, which each man individually satisfies, the incomplete idea of being something is an idea which each thing individually satisfies. Since anything at all by itself satisfies the idea of being something, we’ll say that the idea has universal range. On any particular occasion, one may apply an incomplete idea to anything which satisfies the idea: one may apply the idea of a man to some particular man, the idea of an animal to a particular animal, or the idea of something to a particular thing. On these occasions, the things to which we apply incomplete ideas are their objects. Because the content of an incomplete idea doesn’t specify any particular thing as its object, which thing is its object depends on some intention to apply the idea to
a particular thing (and not to any others). The idea of being something has universal range because it may take any particular thing as its object. The number 1 (a single unit) has universal range. On any particular occasion, we may apply it to any particular thing and that thing will satisfy the idea which is the number 1.

Numbers greater than 1 consist of at least two units. Each unit in any number greater than 1 is an incomplete idea of a thing in general distinct from whatever things which are, on any particular application, the objects of the other units. Each unit still has universal range. But, on any particular application to some objects, the additional content of each unit restricts which things serve as objects for the other units. For the number 2, for instance, each unit has universal range. But on any particular application, the object of one unit will not satisfy the other unit (since it otherwise would be identical to something the idea represents as being distinct from). Given the content of that second unit, whatever satisfies that unit will be distinct from the object of the first unit. So on a particular application of the number 2, the object of each unit will be distinct from the object of the other. This is why the number 2 numbers pairs of things only.

Like traditional units views, the present view requires infinitely many different units. The number 1 consists of a single unit. 2 consists of two additional units. 3 consists of three more units. And so on. Unless there is some reason in principle to deny that there is a collective idea whose components are units from different numbers, there will be infinitely many collective ideas which consist of two units, infinitely many collective ideas which consist of three units, and so on. Does this proliferation of collective ideas of units imply that there are infinitely many number 2s, 3s, and so on? No. The collective ideas which are numbers are well-formed. A unit in the number 3 includes the idea of something distinct from something and some other thing, whereas each unit in the number 2 includes the idea of being distinct from some other thing. Suppose we pair a unit from the number 3 with a unit from the number 2 to form a new collective idea. Such an idea will include the idea of something which is distinct from the object of the other unit; and the other unit will include the idea of something which is distinct from the things a pair of other units represent. But there is only one other unit in the collective idea for its own represented thing to be distinct from it. It represents distinctness from the something represented by an additional unit, one which nowhere appears in the collective idea. This collective idea misfires in what it represents—it is ill-formed. A collective idea of units is well-formed when each unit is such that it represents a thing distinct from all and only
the things represented by the other units. If any unit fails to represent something which is distinct from something represented by any other unit, then a collective idea of units is not well-formed. The natural numbers are well-formed collective ideas.

Here is how numerical idealism would explain counting. Suppose I give you a barrel and ask you to open it and tell me how many things are inside. If there is nothing in there, you observe that there is nothing in the barrel. You form the collective idea of every (mid-sized thing) in the barrel—*nothing*—and say there are zero things in the barrel. If there is a single book in the barrel, you observe the book and form the collective idea of what’s in the barrel, the idea of the book. Then you move to the abstraction of that collective idea, the idea of something. You then say that there is one thing in the barrel. Or if there is a pair of books in the barrel, you instead form the collective idea of the one book and the other. Then you form the abstraction *something and something else*. You say there are two things in the barrel. What explains why you’re in a position to say how many things there are in the barrel after abstracting from the collective idea of what’s in the barrel? The abstraction you grasp in each case is itself the number of things in the barrel.

This view, which is here presented partially and informally, is a variant of the units view. The number 1 is a single unit, the idea of being something. And the components of the collective ideas which are numbers greater than 1 are also units. How does this sort of idealism fare? I develop the view further as I consider various objections.

### 7.7 An Evaluation of Numerical Idealism

In this section, I shall argue that numerical idealism not only overcomes the traditional problems for units views. It also overcomes the kinds of problems facing Fine’s units view. Then I’ll show how numerical idealism overcomes common objections against the view that numbers are mind-dependent objects.

First, numerical idealism resolves the equinumerosity problem and explains why equinumerous collections have the same number of individuals. We can define an abstraction function which takes the collective idea of some individuals and provides its abstraction, in my sense. Suppose we begin with Tom, Dick, and Harry, on the one hand, and Tina, Daria, and Hilda, on the other. There is the collective idea *Tom and Dick and Harry* and another collective idea *Tina and Daria and*
Hilda. The abstraction function essentially replaces each component in a collective idea with a unit, the idea of something which is distinct from whatever things the other units represent. Thus, the abstraction function takes Tom and Dick and Harry and provides the abstraction something, another something, and something else. And the function takes Tina and Daria and Hilda and also provides an abstraction something, another something, and something else. The function takes a collective idea and provides an equinumerous collective idea of units.

What ensures that the collective idea of units is the same in both cases? What generally ensures the identity of a collective idea \( x \) and a collective idea \( y \) is that \( x \) and \( y \) have the same component ideas. And, generally, if the components are not collective ideas themselves, a component of \( x \) is identical with a component of \( y \) when they have the same parts. So we might expect a collective idea of units \( x \) to be identical with a collective idea \( y \) when each unit of one is a unit of the other. The identity of collective ideas of units is more complicated than this, however.

Let’s go through the cases. 0 and 1 are simple. 0 is the idea nothing, the result of abstraction from any empty collection. 1 is the idea of something, the result of abstraction from any collection with a single thing. There is only one idea of nothing and one idea of something. So the result of abstracting from any empty collection is the same and the result of abstracting from any idea of a single thing is the same.

The complications arise for numbers greater than 1. The result of abstracting from any collective idea of individuals is a well-formed collective idea of units. So we need to show that each well-formed collective idea of units is unique. Each unit in a well-formed collective idea includes an iteration of the idea of being something. And each unit in a well-formed collective idea includes the ideas of being distinct from whatever the other units represent. But there is nothing in particular which each unit represents, and so there is nothing in particular each unit represents something as being distinct from. If some well-formed collective idea were not unique, then there should be some difference in what the indiscernable collective ideas of units represent. But there is not. Each unit from a collective idea of units has universal range; and on any particular occasion, the object of each unit will be distinct from the objects of the other units.

The abstraction problem says that abstraction from a set of individuals either provides an equinumerous set of impure units or a singleton set of a pure unit. Units views require numbers greater than 1 to consist of more than a single pure unit. In order for there to be multiple units, then,
each unit must have some feature that distinguishes it from the others. If, in abstraction, we subtract from the ideas of things we count the ideas of what distinguishes them, then abstraction from two or more things will result in a single unit.

Numerical idealism diagnoses this problem as a simple confusion. Abstraction, the subtraction of mental conjuncts, allows one to represent something’s being distinct from something else without representing any particular thing or any particular thing’s particular features. When we abstract to a collection’s number, we abstract to a collective idea of units. Each unit in numbers greater than 1 is a complex idea. It includes not only an idea of being something, but ideas of being distinct from whatever the other units represent. Each unit is pure: no unit includes an idea of having a certain shape, color, size, or whatever. All such ideas are lost in abstraction. But the units themselves are distinct because they point to each other as not representing what the others represent. On any particular occasion, each unit of such a number applies to something which isn’t what any other unit applies to. The distinctness of units arises from the mere distinctness of what they represent. And abstraction allows us to subtract all the parts of each component in a collective idea until we arrive at a collective idea of units which represents mere distinctness.

Does the current view also resolve the plurality problem—i.e., can it explain how the units which comprise each number are both indistinguishable and distinct? There are no units in 0, and only one unit in 1. So the plurality problem arises for numbers 2 and greater. Each unit in such a number contains the idea of being something and the idea of being non-identical to whatever it is the other units are ideas of. Each unit has an identical idea as a part, the idea of being something. And each unit has a part identical in form, the idea of being non-identical to whatever thing the other units are ideas of. In this sense, each unit in a number is indistinguishable from the other units.

Units from a single number are distinct. Each unit represents something distinct from whatever the other units represent. In order for us to apply a unit from a collective idea to something properly, we must apply it to something different from whatever things we apply the other units to. And for ideas $F$ and $G$, if on any particular occasion we can only properly apply $F$ to something which is distinct from whatever we apply $G$ to, then $F$ and $G$ must also be distinct. Each unit is also distinct from other units. It is easy to show that each unit from some number is distinct from the units of every other number. The unit from a number $n$ includes the ideas of being distinct from each of the other $n$
- 1 things represented by the other units. Units from greater numbers include ideas concerning the distinctness of more than \( n \) things and units from lesser numbers include less of these ideas.

There are three problems with Fine’s view, and numerical idealism solves each of them. The first of the problems is that isn’t clear what variable objects are because it isn’t clear what their essential function of having a range of values is. Second, it isn’t clear why zero is a set even though numbers greater than zero do not seem to be tied essentially to sets. And, finally, since Fine only says that variable objects are abstract objects, it is doubtful whether we have any connection to them. I’ll respond to this last problem in the next section. I’ll conclude the present section with responses to the first two problems.

In Fine’s account, the ability of units to take a value on a particular occasion allows them to number the collection of their values on that occasion. If the units \( u^1_2 \) and \( u^2_2 \) take Tom and Jill as their values, respectively, then \( \{u^1_2 \text{ and } u^2_2\} \) numbers \{Tom, Jill\}. And any unit’s universal range is what supposedly accounts for each number \( n \)’s ability to number any \( n \)-membered collection whatsoever.

Since \( u^1_2 \) and \( u^2_2 \) each have a universal range, \( \{u^1_2 \text{ and } u^2_2\} \) numbers any collection \( \{x, y\} \) for which \( x \) and \( y \) are distinct. It is essential to Fine’s units that they have these functions. But it is not clear what it means for a unit to take a value on any particular occasion. And so it isn’t clear what it means for Fine’s units to have a universal range of values, an ability to take anything as a value.

Numerical idealism isn’t subject to these kinds of worries. Instead of “taking a value” no a particular occasion, a unit in my account applies to an individual. And a unit applies to an individual when, first, the individual is identical to itself, and, secondly, when we intend to apply the unit to that individual. The second condition involves an intention on someone’s part to think of a particular individual as “a something.” We apply a unit apiece to Tom and Jill with an intention of thinking of Tom and Jill each as something distinct from the other. Doing so on any particular occasion is an instance of counting Tom and Jill as two things. Of course, 2 numbers Tom and Jill whether we count Tom and Jill or not. Tom and Jill are mind-independently something and something else—and so they satisfy the idea \textit{something and something else} whether we intend to apply the collective idea to them or not.

Any \( x \) and \( y \) which are non-identical satisfy the idea \textit{something and something else}. The universal satisfaction of \textit{something and something else} by any \( x \) and \( y \) which are non-identical is what accounts for the number 2’s numbering any pair of things. And the universal satisfaction of of \textit{something and}
something else by any \(x\) and \(y\) which are non-identical is what allows us to apply the idea to any particular pair of things and truly say of any such pair that it consists of two individuals. Whereas Fine’s account of units leaves their central functions mysterious, numerical idealism explains the central functions of units in terms of the satisfaction of an idea (in the case of numbering) and the intention to apply an idea to the thing or things which satisfy it (in the case of counting).

Why is zero a set? Suppose there is nothing in the barn. The set of what is in the barn, according to Chapter 5, is the collective idea of individuals in the barn. In this case, that idea is the idea of nothing (or none), the empty set. The number of what is in the barn is the collective idea of units which jointly represent how many somethings there are in the barn. In this case, that idea is also the idea of nothing (or none). Zero is the idea nothing, which, according to Chapter 5, is also the empty set.

## 7.8 Defending Idealism

Numerical idealism identifies numbers with mind-dependent objects. But this identification is typically thought to be problematic. I have already surveyed some of Frege’s influential worries against idealist views of number. These include the worry that abstraction does not reliably provide the number of a collection of individuals and also the worry that there would be as many number 1s and 2s as there are ideas of 1 and 2.

Although there are numerous abstractions from a single collective idea of an individual—think of all the ways in which one can subtract any single mental conjunct from any of the components—only one abstraction from a collective idea is the number of the individuals the collective idea concerns. And it is not mysterious what the relationship is between the abstraction and the collective idea or how we abstract from a collective idea to the number of individuals the collective idea concerns. First, take a collection of individuals and form the collective idea of them. Then subtract from that collective idea any of the distinguishing features of each component until all that’s left in the place of each component is the idea of something, along with the ideas of being distinct from whatever other things the other ideas represent. There is a simple diagnosis of those in Frege’s story who abstract to various other kinds of ideas. They aren’t aware that counting, which they do well, is a specific kind of abstraction, which they don’t know much about. They’re ignorant of the underlying mechanics. If
we told those characters simply to count the objects in a collection, they would abstract to a collective idea of units: something (so, one), and something else (so, two), and something else (so, three), and so on, for however many individuals are in the collection. That someone doesn’t abstract to the correct collective idea of units when they are asked to (as in Frege’s example) is no more evidence against Cantorian abstraction than it is against the view that people count with awareness of that counting as an instance of Cantorian abstraction.44

Nor do we need suppose that numerical idealism implies that there are as many different number 1s and 2s as there are people who have the ideas those numbers are. I have characterized ideas as the intentional or representational contents of a mental acts. Mental acts are private to individuals. My grasping of a certain idea involves my mental apparatus, not anyone else’s. But why think mental acts of different persons cannot share the same intentional or representational content? There is an important sense in which what I think when I think something and something else is the very same thing as what you think when you think something and something else. The ideas within numerical idealism are shared ideas of human cognition. Characterizing numerical idealism in this way is not idiosyncratic. According to John Burgess, for example, conceptualism or idealism is the view that “numbers exist, but only as shared human concepts or ideas.”45 He lodges two criticisms against this view. In the rest of this section, I will summarize his criticisms and respond to them.

Burgess offers an analogy by which he hopes to explain why so very few philosophers hold the view that numbers are ideas. He asks us to suppose that there are bigfeet, both undiscovered bigfeet and two known bigfeet, Harry and Harriet. He then immediately distinguishes bigfeet from ideas of bigfeet:

...Harry, Harriet, and other large, hairy, humanoid creatures inhabiting the wilder parts of the Pacific Northwest are very different sorts of things from shared human ideas and concepts, and in particular are very different sorts of things from the ideas and concepts of Harry, of Harriet, and of Bigfoot in general. They differ in absolutely fundamental respects, for instance, in their location in space and time.46

44 Perhaps the criticism never made its way into print because Frege eventually thought of this alternative explanation.
45 Burgess (2003, 24).
46 Burgess (2003, 26).
If the human ideas of bigfeet are located at all, they do not share the locations of bigfeet themselves. How could human ideas of bigfeet inhabit the whereabouts of yet undiscovered bigfeet?

Burgess also notes a particular kind of difference between ideas of bigfeet and (continuing on the supposition that there are some bigfeet) bigfeet themselves:

The creatures differ from the ideas also in respect of how many of them there are. People have ideas of Harry, Harriet, and several more Bigfeet that have allegedly come into contact with human beings; but there are supposed to be, according to the minority view I have ask you to assume for the moment, more Bigfeet than just these: more individuals like Harry and Harriet than there are shared human ideas of individual Bigfeet. The term “Bigfoot” refers to the inhabitants of the wilds of Washington and Oregon, not to the contents of the minds or brains of the cryptozoologists assembled in Scotland. If we wish to refer to the latter, we must use some other expression than the word “Bigfoot,” such as the phrase “the idea of Bigfoot.”

Human ideas of bigfeet cannot be bigfeet, since the latter may outnumber the former. Furthermore, suppose Harriet and her fellow bigfeet perished. In that case, bigfeet and ideas of bigfeet would differ to an even greater extent.

Burgess claims the situation is similar in the case of numbers. Like bigfeet and human ideas of bigfeet, numbers and human ideas of numbers do not have similar spatio-temporal locations. At the very least, there have not always been human ideas of numbers. So were there not numbers before there were animals with sufficiently complex mental lives? The numerical idealist might reasonably grant that there were not numbers. Even if there weren’t ideas of distinct somethings, there would still have been all the particular distinct somethings. Why would there have to have been numbers in addition? We use numbers now to count and measure those particular distinct somethings which once existed. But in that case, we are using presently existing human ideas to count distinct particular things in the past. The numerical idealist who identifies numbers with shared human ideas may bite the bullet here and explain away the initial uneasiness one feels about the notion that numbers came into being.

47 Burgess (2003, 26).
48 Burgess (2003, 27).
But the second objection against idealism is quite serious. Classical mathematics, which I want to endorse,\(^49\) says that there are infinitely many numbers. And, as Burgess says, “surely human beings have formed ideas or concepts of only finitely many of them. There simply are not enough human ideas or concepts for each number to be one.”\(^50\) One may reject classical mathematics, but that is a steep price to pay. What other avenues might an idealist take?

One might, like Geoffrey Hellman, identify numbers with possibilia.\(^51\) Then one might identify numbers with possible ideas, ideas which can be formed of distinct somethings. Although such a position may then provide enough possible ideas to account for numerical reality, new problems arise. What are possible ideas? Are they themselves a sort of abstract object? How do we grasp them? What is the relation between the actual human idea which we use to count pairs of things and the possible idea which is actually the number 2?

Another position, which I’ve not seen anyone defend in print, says that numbers inhabit a mind-dependent realm which pops into existence in its entirety once there is a single cognitive apparatus which exists and is able to grasp at least some portion of it.\(^52\) One could say that the existence of thinking things pushes into existence not only the ideas which are thought, but also those ideas which are thinkable. This doesn’t seem so crazy to me, or at least not any crazier than any other metaphysical view of number. It very well could be true. The challenge is to explain why thinking itself has the amazing creative power to produce simultaneously all the thinkables. Otherwise, the view does seem like an ad hoc solution to the present problem.

Lastly, one might identify numbers with ideas in an infinite mind which supplies those ideas as the raw materials of thought for all other thinking beings. This is Cantor’s view (and also the view of many others through the ages).\(^53\) And it has two explanatory benefits. The proposed nature of this infinite mind helps explain not only the existence of all numbers (for there would be enough ideas with which to identify numbers) but our epistemic access to numbers, too. Such a view might

---

\(^49\)See Chapter 5, Section 1.

\(^50\)Burgess (2003, 28).

\(^51\)Hellman (1989).

\(^52\)My brother, Brandon Warmke, has defended such a view in conversation—in Facebook Chat. And although it seems to have been one of his favored views and not merely an ad hoc response to the present problem, I do wonder whether he wanted to elicit a reaction from me.

\(^53\)See Dauben (1990, Ch. 6 and 236–239) and Hallett (1986, 9–11, 35–36).
say that the infinite, omnipotent mind holds ideas in being and endows them with causal powers by channeling its own causal efficacy.\textsuperscript{54} If such a view were true, we could partially explain our epistemic access to numbers by appealing to the causal connection we have to them. The explanation would only be partial, however, because the nature of that causal connection would likely remain mysterious.

\textsuperscript{54}See Chapter 4, Sections 15 and 16.
BIBLIOGRAPHY


