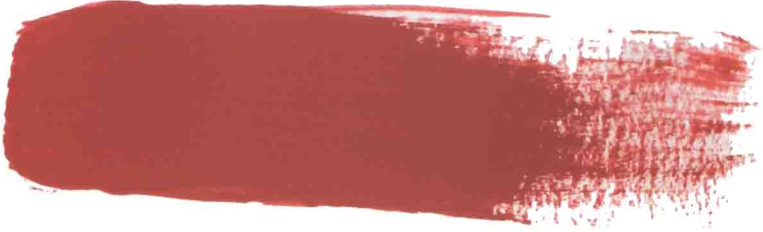
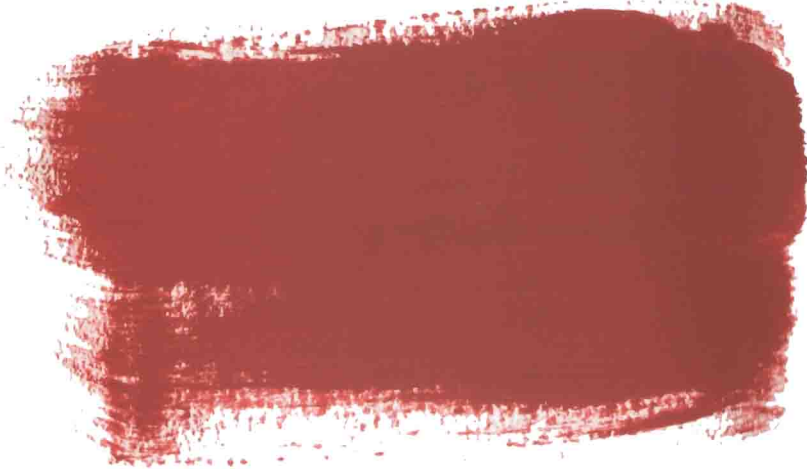




Mere Possibilities



METAPHYSICAL FOUNDATIONS
OF MODAL SEMANTICS



Robert Stalnaker

Mere Possibilities



METAPHYSICAL FOUNDATIONS OF MODAL SEMANTICS

Robert Stalnaker

It seems reasonable to believe that there might have existed things other than those that in fact exist, or have existed. But how should we understand such claims? Standard semantic theories exploit the Leibnizian metaphor of a set of all possible worlds: a proposition might or must be true if it is true in some or all possible worlds. The actualist, who believes that nothing exists except what actually exists, prefers to talk of possible states of the world, or of ways that a world might be. But even the actualist still faces the problem of explaining what we are talking about when we talk about the domains of other possible worlds. In *Mere Possibilities*, Robert Stalnaker develops a framework for clarifying this problem, and explores a number of actualist strategies for solving it.

Some philosophers have hypothesized a realm of individual essences that stand as proxies for all merely possible beings. Others have argued that we are committed to the

Stalemaker

Mere Possibilities



PRINCETON

{ Mere Possibilities }

Carl G. Hempel Lecture Series

{ Mere Possibilities }

Metaphysical Foundations
of Modal Semantics

Robert Stalnaker

Princeton University Press
Princeton and Oxford

Copyright © 2012 by Princeton University Press

Published by Princeton University Press, 41 William Street, Princeton,
New Jersey 08540

In the United Kingdom: Princeton University Press, 6 Oxford Street,
Woodstock, Oxfordshire OX20 1TW

press.princeton.edu

All Rights Reserved

ISBN 978-0-691-14712-3

Library of Congress Control Number: 2011939335

British Library Cataloging-in-Publication Data is available

This book has been composed in Minion Pro with Bauer Bodoni Std. 2
Display Black

Printed on acid-free paper. ∞

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

To my students,
who have taught me so much philosophy,
among other things.

{ Contents }

Preface ix

1

On What There Isn't (But Might Have Been) 1

2

Merely Possible Possible Worlds 22

3

What Is Haecceitism, and Is It True? 52

4

Disentangling Semantics from Metaphysics 89

5

Modal Realism, Modal Rationalism,
Modal Naturalism 126

Appendix A

Modeling Contingently Existing Propositions 136

Appendix B

Propositional Functions and Properties 139

Appendix C

A Model for a Mighty Language 149

Appendix D

Counterpart Semantics for the Cheap Haecceitist 154

References 157

Index 161

{ Preface }

I have been thinking about possible worlds and making use of the apparatus of possible-worlds semantics since I took a seminar taught by Saul Kripke in my last year of graduate school at Princeton in 1964–65. In my early work that used that framework, on the semantics for conditionals, the representation of propositional content, and the dynamics of discourse, I didn't worry much about the metaphysical questions—about what possible worlds and merely possible individuals are, and whether it is legitimate to take them seriously. The idea seemed clarifying, and the semantic framework seemed to yield results, and that was good enough for one who prided himself on his lack of an ontological conscience. But I was puzzled (and ultimately chastened) by a remark by Larry Powers in an insightful commentary on an early paper of mine on propositions: “The whole idea of possible worlds (perhaps laid out in space like raisins in a pudding) seems ludicrous.”¹ At the time, it had not occurred to me that one might think of possible worlds as parallel universes, but I came to see that if one is to reject this literal-minded interpretation of the term (which I soon learned was defended by David Lewis), one needs to say something about what these things are. I tried to do this in a paper, “Possible Worlds,” first published in 1976, but that paper is silent about a further question about merely possible individuals: How, on an actualist interpretation of possible worlds as ways a world might be, is one to account for the possibility that there be individuals other than those that actually exist? That is the main focus of this book.

¹ Powers 1976, 95.

Responding to the problem led me into a tangle of metaphysical issues. I have always been a reluctant metaphysician—one who acknowledges that metaphysical questions cannot be avoided but who continues to be puzzled about their nature. While my primary aim in this book is to say something about the substantive questions of modal metaphysics, I also have a secondary aim: to get clearer about metaphilosophical questions about the nature of metaphysics and about the relation between semantic and substantive philosophical questions. For the most part, I don't address the meta-questions directly in the book. The best way to approach them, I think, is to focus on first-order metaphysical questions, keeping an eye, and occasionally commenting, on what one is doing as one is doing it.

While I have, and express in the book, a substantive view about the metaphysics of mere possibilities, I also try to develop a common framework for representing alternative metaphysical pictures and to make as coherent as I can the metaphysical pictures that I ultimately want to reject. I think this helps clarify, by contrast, the picture I want to defend, but it also tends to sharpen the puzzle about the nature of metaphysical theses. How do we choose between formally coherent alternative metaphysical theories? I don't have a complete answer to this question to offer, but I hope what I say will be relevant to it.

This project began with an informal talk, some years ago, to the Arché group at the University of St. Andrews. The talk grew into a paper that eventually became chapter 1 of this book. Chapter 2 overlaps with a second talk given at a conference on modality at St. Andrews and published in the proceedings of that conference.² The invitation to give the Hempel lectures at Princeton University provided the occasion for further development of the ideas. The first three chapters were based on those lectures, given in May 2009. A

² Stalnaker 2009.

month later I gave the Pufendorf lectures at the University of Lund, adding a fourth lecture to those given at Princeton. An expansion of this lecture became chapter 4. I am grateful to Arché and the philosophy departments at Princeton and Lund for giving me the opportunity to develop and present these ideas and to the audiences at these occasions for stimulating and helpful discussion.

Thanks to Agustín Rayo, Bob Hale, and Damien Rochford, who read a complete draft of the manuscript and gave me very helpful comments that led to what I hope are improvements. Thanks to Aviv Hoffmann and Delia Fara for discussion and comments. Thanks to my editor, Rob Tempio, for his support and advice. Thanks also to an anonymous referee for Princeton University Press, who gave me insightful comments that led to significant revisions. For editorial help at the late stage of preparation of the manuscript, thanks to copy editor Jennifer Backer and Damien Rochford.

I was particularly pleased to have the opportunity to give lectures that honored C. G. Hempel, who was my teacher and supervisor at Princeton, as well as a philosopher whose writings helped draw me into philosophy years before I came to know him. He has long been a role model for me for his clarity of mind, his generosity, and his integrity.

The audience for the lectures at Princeton that honored one of my teachers included two others who had been among my graduate teachers at Princeton, Paul Benacerraf and Gil Harman, still hanging around the place more than forty years later. It also included four philosophers I had taught, former students in the graduate program at MIT who are now on the faculty at Princeton: Adam Elga, Delia Fara, Liz Harman, and Sarah McGrath. The occasion led me to reflect on the relationship between graduate students and their teachers. I felt like a link in a chain that goes back to the heyday of logical empiricism and forward long into the future. I learned a lot and was profoundly influenced by my undergraduate and graduate teachers, but when I became a teacher myself I

learned that the impact goes the other way as well. Interaction with a really excellent group of graduate students in philosophy at MIT over the twenty-three years I have been there has challenged and inspired me, helped keep me open to new ideas, and influenced the direction of my work. I take pride in their accomplishments, but mainly I want to thank them for their contributions to my understanding of philosophy. This book is dedicated to those students and former students.

Cambridge, MA

January 2011

{ Mere Possibilities }

{ 1 }

On What There Isn't (But Might Have Been)

The problem of ontology, Quine told us in his classic essay “On what there is,”¹ can be put in a simple question, “what is there?” and answered in a word: “everything.” My question should be equally simple, and its answer should follow from Quine’s: there is nothing that isn’t. But of course as Quine went on to say, the problem gets harder when one tries to be more specific about what there is and what there isn’t. Quine’s concern was mainly with the problem of expressing disagreement about ontology—if I believe there are more things in Heaven and Earth than are dreamt of in your philosophy, how can you talk about what it is that I believe in, but you do not? But even when we agree about what there is, we may want to acknowledge that things might have been different—not just that things might have been differently arranged but that there might have been different things than there actually are. If we ask not just “what is there” but “what might there have been,” the answer “everything” does not seem sufficiently inclusive. But what else is there that might be included?

The problem is sufficiently daunting to have driven many philosophers, in different ways, to deny there could have been anything other than what in fact exists, or that anything that exists could have failed to exist. (Three examples of philosophers who develop this idea in very different ways: Wittgenstein of the *Tractatus*, David

¹ Quine 1948.

Lewis, and Timothy Williamson.) Others have hypothesized actual surrogates for the nonexistent things—individual essences that are themselves necessary existents and that correspond one-to-one with all the “things” (as we are inclined to put it) that might exist.² Still others think that because taking modality seriously forces us to such metaphysical extravagance, we should reject modal discourse as anything more than a *façon de parler*. But I think modal concepts are central to our understanding of the world—the actual world—and that understanding them should not require extravagant metaphysical commitments. My aim in this book is to sketch a framework that allows us to avoid extravagant metaphysical commitments and that is also compatible with intuitively natural beliefs about the way things might have been.

There are some philosophers who want to take modality seriously, and seek a theoretical account of modal discourse, but who think that we cannot take possible-worlds semantics, as an account of modality, seriously without making extravagant metaphysical commitments. Christopher Peacocke, for example, holds that “it is an unstable, indeed incoherent, position to think that you can at the same time use the Kripke-style semantics in the metalanguage to give absolute truth-conditions for modal sentences, count . . . [the proposition that there could have been something that doesn’t actually exist] as true, yet avoid commitment to the existence of nonactual objects.”³ But I want to defend the metaphysical innocence not only of modal concepts but also of a theoretical account of them in terms of possible worlds. Whether my construal of possible-worlds semantics counts as a realistic one or not is open to debate, and I will concede that on one of the several ways of construing the term “possible world,” the possible worlds posited by these semantic models are artifacts of the model and not entities

² This is the response to the problem developed and defended by Alvin Plantinga. See the papers collected in Plantinga 2003.

³ Peacocke 2002, 121.

whose existence is affirmed. But I will argue that on another way of understanding the term, we can affirm the existence of possible worlds, as well as the claim that the semantic theory provides "absolute truth conditions for modal sentences" and "avoids commitment to the existence of nonactual objects."

Here is my plan for this chapter: I will start, in section 1, with some preliminary methodological remarks—about the aim and value of reduction in philosophical analysis, about thinking of the evaluation of philosophical theses in terms of costs and benefits, and about the contrast between realistic and anti-realistic accounts of a philosophical theory. In section 2, I will say what I take possible worlds to be, and what, from the perspective of this account of possible worlds, the problem is about merely possible individuals. Possible worlds, on the account I want to defend, are (to a first approximation) *properties*, and the main point I want to make in this section is that properties (and so possible worlds) are not *representations*. In section 3, I take an extended look at some examples of properties that are simpler and easier to think about than possible worlds but that share some of the features of possible worlds, construed as properties. In this section and section 4, I will use the analogy I develop to motivate what I hope is a metaphysically innocent account of the domains of other possible worlds.

The view I will be defending is committed to making sense of the contingent existence of individuals and properties, of propositions, and even of possible worlds themselves. I will conclude, in section 4, by sketching a problem that an account of this kind faces, a problem that I will respond to in chapter 2.

1. Methodological Preliminaries

According to John Divers in his useful survey of the range of alternative philosophical accounts of possible worlds, "the primary question of conceptual application of the species of AR [actualist

realism] is whether any affords a thoroughly non-modal analysis of the family of modal and intensional concepts.”⁴ Divers acknowledges that “the proponents of AR typically do not claim that the favored version of AR affords thoroughly nonmodal analysis of the modal concepts,”⁵ but he seems to assume that it would be a benefit (in the cost-benefit evaluation of the general view) if it did provide such an analysis. But my view is that if an account of modality were to meet this condition, that would be a sure sign that it was on the wrong track. Necessity and possibility are fundamental concepts, like truth and existence. What would you say to a philosopher who was seeking a thoroughly nonexistent analysis of quantificational concepts, or a thoroughly non-alethic analysis of truth, and related concepts? It is not that philosophers have not proposed such analyses (substitutional quantification, truth as warranted assertability or as what ideal believers will believe at the end of inquiry, for example). But even if an analysis of this kind were to be extensionally correct, at least according to someone’s philosophical theory, it would only blur the distinction between semantic analysis and a substantive metaphysical thesis about what exists or what is true. Consider the nominalist who defines existence as *having spatio-temporal location*. Platonists will agree that if that is what you mean by “exist,” then numbers, sets, and properties do not exist. They will need to find alternative means of describing their ontological beliefs.

I do not want to suggest that one can distinguish, on some pre-theoretical a priori ground, which concepts are fit subjects for some kind of reductive analysis. It may be a contentious philosophical question, not only how to answer substantive questions but also which questions are substantive and which are semantic. So, for example, I am inclined to think not only that what is *actual* coincides with what exists but that this is because “actual” just means

⁴ Divers 2002, 181.

⁵ Ibid., 301n.

(more or less) real, or existent. The modal realist disagrees, and he might complain that by understanding "actual" in this way, I am blurring the line between metaphysical and terminological questions. I agree that my disagreement with the modal realist is a mix of semantic disagreement and disagreement about what there is in the world, and that to be clear, it is important to try to separate semantic from substantive questions, but it is not always easy to do so.⁶ I will discuss this issue in more detail and make some claims about how the two kinds of issues should be separated in chapter 4.

I have alluded to the cost-benefit, reflective equilibrium methodology that Lewis articulated and made fashionable, but I have my reservations about this way of thinking about the way philosophical alternatives are evaluated. This picture may be fine if it is taken simply as a reminder that in philosophy, as in science, political theory, or any other enterprise, everything is potentially criticizable; there are no absolute unquestionable dogmas. One should add that even judgments about what is a cost and what a benefit might be a proper subject of debate. But beyond the bland truism, the reflective equilibrium method does not offer much guidance. Even though anything might be epistemically relevant to anything else, one important task, in deciding between alternative philosophical views, is to isolate considerations of different kinds. There may be no absolutely neutral conceptual standpoint, but it is a virtue of a theoretical account of some concept or family of concepts (a benefit in the cost-benefit analysis) if it is able to fashion some tools that manage to remain neutral on issues in dispute—to provide resources to formulate alternative substantive views as coherently as possible. A more neutral account (of truth, existence, properties and relations, modality) may seem disappointing (it would be nice to have an account of truth that gave us a lot of information about

⁶For an excellent discussion of some of the problems of sorting out semantic from metaphysical questions when discussing fundamental ontology, see Lewis 1990.

what is true), but I think we should be suspicious of an account of modality that tells us too much about what there is or about what there might have been.

Consider this parody of the cost-benefit methodology, run amok: X says, "I have a beautiful, austere, and crystal clear theory of properties: they are just sets—no more and no less. The relation between a property and its exemplifications is just the relation between a set and its members." Y responds: "It is a beautiful theory, I agree, but unfortunately it is false—there are many obvious counterexamples. We don't need to consider exotic examples like renates and cor-dates. Consider any two uninstantiated properties like being a talking donkey and being a philosophizing cat [two of David Lewis's favorite examples]. It follows from your view that these two properties are one, which is obviously wrong." X replies: "Some who like my theory—perhaps a Quinean—would reply by rejecting your intuition that there are distinct properties here. But I am a commonsensical chap [as David Lewis liked to describe himself], and I agree with you that the properties you have described are distinct. Nevertheless, I am reluctant to give up my beautiful theory, since its benefits are great. I prefer to give up instead the belief that there are no talking donkeys and no philosophizing cats."

There is much more that X needs to say, for example, about how these ontological hypotheses are to be reconciled with apparent evidence to the contrary, but however X goes on, I think most of us will find this response suspicious, not just because the benefit of the beautiful theory is outweighed by the cost of the ugly facts, but because there is something suspect about using this kind of theoretical virtue to reach this kind of conclusion about what there is. (This is a parody, but it can be argued that my story of X and Y is just an uncharitable spin on the kind of consideration that actually motivated Lewis's modal realism. Lewis does hold this theory of properties, and getting the identity condition for properties right is a prime motivation for the hypothesis of a plurality of worlds. And

it is the identification of properties with sets that rules out, for him, an actualist account of possible worlds.)

One final methodological remark before getting down to business: it is common to distinguish between “ontologically serious” applications of modal semantics and purely instrumental uses; the latter includes mathematical uses (for example, the construction of models to show the satisfiability, in a technical sense, of certain sets of sentences of a formal language) and heuristic uses that treat possible worlds discourse as “a vivid shorthand for sentences containing modal operators.”⁷ It is often suggested that if we are to take possible-worlds semantics to be a theory that contributes to a project of philosophical explanation of modality, then we must specify a particular model—the intended model of metaphysical possibility. Jon Barwise and John Perry wrote, in criticizing possible-worlds semantics:

If the model-theoretic structures of possible worlds semantics, the ones that include a set of all possible worlds, are supposed to be a model of something, say super-reality, under some correspondence or other, then there ought to be *one* that is an intended or standard model, *the* one that really corresponds to super-reality.⁸

I think this contrast is oversimplified. Taking possible-worlds semantics seriously as an explanatory account need not require the belief that there is one intended model any more than taking quantification theory, and its semantics, seriously requires the belief that *the* intended interpretation is in terms of a single domain of absolutely everything. There is controversy about whether it makes sense to quantify over absolutely everything, but whether it does or not, I think all should agree that we can take quantification theory as more than a mathematical tool or a heuristic device even while

⁷ Sider 2002, 280.

⁸ Barwise and Perry 1985, 120.

rejecting the idea of an absolute domain. And whether or not it makes sense to talk of an absolute, context-independent domain of all possible worlds, it is useful to separate the project of clarifying the framework for doing modal metaphysics from the project of saying, within that framework, what is really necessary and possible. (Just as Quine distinguished the project of getting clear about what ontological commitment is and how it is to be represented from the project of stating what one's ontological commitments are.) I agree that if we are to take possible-worlds semantics seriously, we must say something about the *kind* of thing that a possible world is and justify the claim that it is reasonable to think that there are such things as possible worlds. But my aim will be to vindicate the possible-worlds theory while making minimal commitments about substantive metaphysical questions, for example, about whether there are things, or properties, that exist only contingently, whether there are individual essences that are irreducible to qualitative properties, whether there could be distinct but qualitatively indiscernible worlds. In the balance of costs and benefits, I give positive weight to this kind of neutrality.

2. What Are Possible Worlds?

So what, on my view, are possible worlds (in the sense in which it is reasonable to say that there is a plurality of such things)? I take them to be *properties*—ways a world might be. Of course this leaves a lot open, since there are many different accounts of what properties are, but I take the significance of the categorization of possible states of the world as properties to be that it implies at least these two things. First, a possible world is the kind of thing that is, or can be, instantiated or exemplified. An actualist needs the distinction between existing and being exemplified in order to be able explain the sense in which a merely possible world exists (a property the world might have had exists) and the sense in which it does not (no

world that is that way exists). But second—and this is the point I want to emphasize—if possible worlds are properties, they are not *representations*—not mental or linguistic entities. So the account of possible worlds I will defend rejects what Lewis calls linguistic ersatzism, as well as the other forms of ersatzism that Lewis considers, all of which treat possible worlds as representations. The significance of this point is that possible worlds are not the kind of thing that faces a problem of intentionality. About a representation (a name, a predicate, a picture, a scale model, a sentence) we can intelligibly ask, what is it about it in virtue of which it represents what it represents? (The following, for example, is a perfectly reasonable question: “What is it about the inscription or vocable ‘triangular,’ as it is used in a certain linguistic community, that makes it a word for the property of being triangular, rather than for the property of being square?”) David Lewis, taking the actualist to be giving some kind of representational account of possible worlds, asks: What is it about a world in which there are talking donkeys that makes it a world in which there are talking donkeys, rather than a world in which there are philosophizing cats? But if possible worlds are properties, this is like the question, what is it about the property of being triangular that makes it *that* property rather than the property of being square? This I take to be an unintelligible question.

The assumption that possible worlds are representations is widespread. Brad Skow, for example, gives voice to the following remark: “Possible worlds are *representations*. All theories of possible worlds agree about this.”⁹ I can say with confidence that the second of these two statements is false, since my own account of possible worlds rejects the first statement, and I don’t think I am alone in rejecting the idea that possible states of the world are representations.

⁹Skow 2008, 103. It is not clear that Skow endorses this remark, since it is attributed to a critic of a point that he is making. But his response to the critic does not reject the claim.

But Skow's claim has an appearance of plausibility that I think rests on the fact that while most philosophers reject David Lewis's modal realism, most have accepted his way of framing the debate about possible worlds. A theorist of possible worlds (many follow Lewis in assuming) is either a modal realist or a believer in ersatz substitutes for the possible worlds that the modal realist believes in. But is it appropriate to describe a property of individuals as an ersatz individual? Is, for example, the property of being a king an ersatz king? Does the property of being a king *represent* something as being a king? What does it represent as being a king? One can *use* properties to represent: the colors, red and blue, for example, are used to represent Republican and Democratic voting patterns, respectively. (It is not just the predicates "red" and "blue" that do the representing, as in the expressions "red state" and "blue state." One also uses the colors themselves, on maps, to do the representing.) Properties such as color properties might be used to represent themselves, as when one colors a part of a scale model of something red in order to represent that the corresponding part of the thing being modeled is red. It might be perverse, but one could also use different colors for this purpose—red to represent blue. So a color might represent a color; nevertheless, the relation between a property and what exemplifies it is not itself a representational relation. By painting the wall blue one does not thereby represent the wall as being blue, nor does the wall itself represent itself as being blue simply by being blue.

Why does it matter that the relation between a property and its exemplifications is not a representational relation? It matters because if one thinks of this relation, or the relation between a proposition and the world in virtue of which the proposition is true or false, as a special case of a representational relation—a particularly intimate one—then one creates the illusion of a problem. When properties or other things are used to represent, one explains the representation relation in terms of the intentions of the users. But

properties and propositions are thought of as mind-independent objects that are *intrinsic* representations: they represent without our help; how do they do it? Jeffrey King characterizes the classical view of propositions as the view that propositions are “eternal abstract entities that *by their very nature and independently of all minds and languages* represent the world as being a certain way and so have truth conditions.”¹⁰ Nothing could do this, which is why King rejects the classical view. One central problem for any theory of propositions, King argues, is to explain their capacity to “represent how the world is,”¹¹ to explain “how propositions have truth conditions.”¹² But on the account of propositions I will defend, propositions *are* truth conditions. What needs to be explained is how things that express propositions—that represent the world as being some way—can express the propositions that they express. But in giving a theory of possible worlds and propositions themselves, we are not addressing this question.

If one tries to say just a little about what properties, in general, are, it becomes clear the extent to which, in classifying possible worlds as properties, we are not explaining modal notions in terms of something more basic. I take the notions of property and relation to be themselves modal notions. Properties are to be understood in terms of what it would be for them to be exemplified, which means we understand what a particular property is in terms of a range of possible situations in which it would be exemplified. But possible situations, we are saying, are themselves properties—ways a situation, or a world, might be. It is not reduction but regimentation that the possible-worlds framework provides—a procedure for representing modal discourse, using primitive modal notions, in a way that helps reveal its structure.

¹⁰ This quotation is from a handout of a talk, but the general view is expounded in King 2007.

¹¹ King 2007, 3–4.

¹² Ibid., 58.

What are possible worlds properties *of*? They are properties of the total universe. One may question whether there is such a thing as the total universe to be what has these properties,¹³ but I will assume that one can intelligibly speak of a universe that is (in the sense of “exemplifies”) a way things might be. (If there is no such entity, perhaps we can speak of possible states of the world as being exemplified, or not, but not *by* anything.)

Possible worlds, on the actualist construal, are usually said to be *complete* or *maximally specific* in some sense. The idea seems to be that they are properties that are as specific as the things that might exemplify them, but it is not easy to say exactly what this means. There are properties that are defined in terms of their exemplifications (like the property of being identical to Osama bin Laden), but of course there is only one possible state of the total universe that is exemplified, and so the others cannot be defined in terms of the universes that exemplify them. What is it for a property that is not so defined to be as specific as what would exemplify it?

Rather than trying to explain what this might mean, I will define maximality in a different way: a possible state of the world must be maximal in the sense that it decides every proposition. But propositions (in the possible-worlds theory) are identified with sets of possible worlds (or equivalently, functions from possible worlds to truth values), and on this account, the claim that possible worlds are maximal puts no constraints on the character of the worlds. If one explains propositions independently of possible worlds (perhaps the propositions are all the properties that are either exemplified or not by the total universe), then we would have an account of what it is for a world to be maximal, at least relative to the domain of propositions. But one might be suspicious of an absolutely complete domain of all propositions. What matters for the applications of possible-worlds semantics is that the possible states of the

¹³ Robert Adams raised this question, in correspondence, in 1974 in response to my original paper on possible worlds.

world be maximal with respect to all questions that are of concern in the application at hand. I prefer to think of the worlds not as the *points* in logical space but as the cells of a relatively fine-grained partition of logical space—a partition that makes all of the distinctions we need. If the partition is fine enough for the purposes at hand, then we can understand the propositions as sets of the partition cells. We do not thereby foreclose the possibility that in some other context, one might cut the space up more finely. The question of whether there is an absolutely finest partition or whether the space is best understood as an atomless algebra, rather than a set of points, is a controversial metaphysical question we can set aside: taking possible-worlds semantics seriously does not require a commitment to an interpretation in which the possible worlds are absolutely specific, in some metaphysical sense.¹⁴

I have been suggesting that possible-worlds semantics need assume only that possible states of the world are as specific as is needed for the purposes at hand, but where the purposes at hand involve understanding talk about what might exist, but does not, we have a problem. The problem is that it seems that in this case, our purposes may require that we carve up logical space more finely than we have resources for. Since we are actualists, we have only the resources that the actual world provides for representing possibilities. We can represent a purely existential possibility (for example, that there is a purple cow) if we can understand the property of being a world in which there is a purple cow.¹⁵ But we understand a property in terms of what it would be for it to be instantiated, and

¹⁴ Cf. Saul Kripke: a “‘counterfactual situation’ could be thought of as a mini-world or a ministate, restricted to features of the world relevant to the problem at hand” (1980, 18). See also Stalnaker 1986, where I distinguished internal from metaphysical completeness (unconsciously echoing Hilary Putnam’s terminology for two kinds of realism) and argued that possible-worlds semantics was committed only to the former.

¹⁵ It is controversial whether we can understand the possibility of a purple cow; Peter van Inwagen (1998) has suggested that such a beast may be impossible.

this general property would be instantiated only if a more specific property, being a world in which a particular x exists, where x is a purple cow. We have a problem if we want to say that while there might have been purple cows, there are no particular things that might have been purple cows.

I am going to approach this problem indirectly by looking at some examples of properties that are exemplified by things that are less grand than total universes but that illustrate some of the problematic features of such properties.

3. Containment Properties

I want to consider a range of properties that an envelope (for example, one of those large envelopes that are recycled in the campus mail) might have, properties that concern what is inside the envelope. Start with these three examples:

- (1) the property of containing three sheets of blank white paper, size A4
- (2) the property of containing a reprint of a critical notice, published in *Mind*, of David Lewis's *On the Plurality of Worlds*
- (3) the property of containing two photocopies of a handwritten letter from Ludwig Wittgenstein to Saul Kripke

Call these *generic containment properties*. One might also define *specific containment properties*, such as:

- (4) containing three *particular* sheets of blank, white paper, size A4 (in a particular order)
- (5) containing *this* reprint of a critical notice, published in *Mind*, of David Lewis's *On the Plurality of Worlds*

(4) and (5) might be construed in different ways. One might mean something like this by (4): *containing exactly a b and c, in that order* (which are in fact sheets of blank white paper, size A4); alternatively, one might mean containing *exactly three sheets of blank white paper, size A4, namely a b and c (in that order)*. (The difference is that on the first understanding, the three specific items might have the property in a possible world in which they are not blank white sheets of paper, while on the second they must be.) I will understand specific containment properties in the second way.

For every generic containment property that is instantiated, there is a corresponding specific containment property that is instantiated by the same thing. We could define this correspondence relation; it would be a second-order binary relation that relates two properties.

Now Saul Kripke (SK) was about twelve years old when Ludwig Wittgenstein (LW) died. We know that Kripke was a precocious child, but I am going to assume that these two philosophers never exchanged letters and thus that there are no photocopies (or things that might have been photocopies) of a handwritten letter from Wittgenstein to Kripke. If this is right, then (it seems reasonable to assume) there will be no specific properties corresponding to our third example of a generic containment property. Still, the general claim we made about correspondence still holds: (3) would be exemplified by the envelope only if a corresponding specific property were exemplified by that envelope.

In terms of this second-order correspondence relation, we can define a second-order property of properties—being a specific containment property corresponding to the generic containment property (3). This is an uninstantiated second-order property but a perfectly good property nonetheless.

One can define properties that are partly specific, partly generic, such as:

(6) containing a certain specific sheet of paper, plus two others (all blank, white, size A4)

One can define negative and disjunctive containment properties:

(7) *not* containing a reprint of a critical notice, published in *Mind*, of David Lewis's *On the Plurality of Worlds*

(8) containing either three sheets of blank white paper, size A4, or a certain specific reprint of a critical notice, published in *Mind*, of David Lewis's *On the Plurality of Worlds*

And one can define additional second-order relations, for example, a *permutation* relation that might hold between two specific containment properties: say that two specific containment properties are *permutations* of each other if they involve the same specific objects, but in a different order.

So suppose we had a specific property corresponding to (3)—the property of containing two specific photocopies of a letter from LW to SK. Then there would be a different specific property that permutes these two specific photocopies. Of course there are no specific properties of that kind, since there are no letters from LW to SK, but there is still no problem with the second-order relation.

This game could go on, but it is time to connect our exercise back to possible worlds. Before doing this, let me point out just one general fact about *negative* generic and specific properties. We noted that if a generic (positive) containment property is exemplified, then some corresponding specific property must be exemplified (by the same thing). So if there are no specific properties corresponding to a generic property, it follows that the generic property is un-instantiated. It is also the case that if a *negative* generic property is exemplified by something, then *every* corresponding negative specific property is exemplified by that thing. For example, if the envelope does not contain three blank white sheets of paper, size A4,

then for every a b and c the envelope does not contain three blank white sheets of paper, size A4, which are a b and c , in that order.

Since there is no possibly instantiated specific property of the form *containing x and y , which are photocopies of a handwritten letter from LW to SK*, there also are no negative specific properties of this kind. But we can see that since the envelope *actually* has the negative generic property, not containing two photocopies of a handwritten letter from LW to SK, if there were a specific negative property of this kind, it would be a property that it seems would be exemplified *in the actual world*. The point is (if I may put it in this loose way) that there are merely possible properties (such as the specific negative properties that would exist if there were any photocopies of handwritten letters from LW to SK) that are *actually* instantiated. We will return to this point.

Possible worlds, we said, are properties, and I hope the way that they are like containment properties is clear. Worlds, like envelopes, have things in them, and they might have contained things other than those they in fact contain. A possible (state of the) world is like a mixed generic/specific containment property. A counterfactual world might be specified as one containing a certain specific thing (Saul Kripke, for example) and a thing of a certain kind that is not any actual thing (for example, SK's seventh son). If the property of being a world containing SK and his seventh son were exemplified, then there would be a more specific property that would also be exemplified (a property of the form *containing SK and x , x being the seventh son of SK*). There are, it seems reasonable to believe, no persons who might have been SK's seventh son, or anything that might have been a person who was SK's seventh son, and so no properties of this form that might be exemplified. But we can still generalize, using second-order properties and relations, about properties of this kind that involve specific individuals.¹⁶

¹⁶ In defending the contingency of properties and propositions, I am following Kit Fine, who has long argued for this. See the postscript in Prior and Fine 1977. The

4. Kripke's Dice: Let Me Count the Ways

I will use an example Kripke used, slightly modified, to expand on the point that one can use second-order properties and relations to talk about the possibility of specific properties that do not in fact exist. Kripke, to underline the modest and commonsensical character of his conception of possible worlds, and to help dissolve what he regarded as a pseudo-problem about the identification of individuals across possible worlds, asked us to consider a simple school probability exercise—a problem about a pair of dice and the thirty-six possible ways that they might have landed. Kripke tells us that “one of these miniworlds—the one that corresponds to the way the dice in fact come up—is the ‘actual world.’”¹⁷ So Kripke is assuming (fictively) that we are talking about an actual pair of dice, which he labels die A and die B. His main point was that it would be silly to ask, about the possibility in which A lands 6 and B 5, how we know that it is A, rather than B, that was the 6. But suppose our dice are a merely possible, generic pair of dice. There is a possible state of the world in which two such dice are thrown, one lands 6 and the other 5, but there is not a different state or property in

account I want to defend is also very close to the view developed in Adams 1981, but there are some differences between our views. Adams distinguishes more sharply than I would between qualitative properties and properties that are ontologically dependent on particular individuals, and he seems to be assuming that while properties of the latter kind may exist contingently, purely qualitative properties will be necessary existents. Perhaps there are some very abstract properties that exist necessarily, but I would argue that qualitative properties such as color and shape, like the things that exemplify them, exist only contingently. But my most important disagreement with Adams is that he holds that a metaphysical view that accepts the contingency of propositions and possible states of the world requires, or at least motivates, a serious modification of modal logic he describes as “metaphysically satisfying though formally inconvenient” (1981, 29). I will argue in chapter 4 that the formal inconvenience is both greater than Adams suggests, and unnecessary; the standard logic and formal semantics can, I think, be reconciled with the austere and satisfying metaphysics.

¹⁷ Kripke 1980, 16.

which two such dice are thrown, one lands 5 and the other 6. Of course we might add some detail to distinguish the two dice: we might, for example, stipulate that one has a scratch on the face with one spot, while the other does not. Then we could distinguish the possible state in which the one with the scratch lands 6 (the other 5) from the situation where the one with the scratch lands 5 (the other 6). But suppose there is no such detail. How do we distinguish the 6-5 situation from the different 5-6 situation? What do we mean when we call one of the dice A and the other B? (That one is A and the other B is not a fact about the possible states that can be used to distinguish them. The A and the B are our labels for describing the situation.) Perhaps we should say, in the generic case, that there are really just twenty-one possible states of the dice but that if one of them had been realized, then there would have been thirty-six possible states of the specific dice that would then have existed. We cannot distinguish specific die A from specific die B, from the perspective of the actual world, where neither exists, but we can talk, in a general way, about specific properties of the form *A lands 5, and B 6*, and we use the second-order permutation relation to talk about pairs of specific properties, both of this form, but with the A and B reversed. We need to talk, in a general way, about the possible specific properties in order to represent facts about the generic situation, such as the fact that in the possible situation in which one lands 5 and the other 6, it is also true of the one that landed 6 that it might have landed 5 while the one that landed 5 landed 6.

5. The Problem of Iterated Modality

Let me conclude this chapter by summarizing the general idea and pointing to a problem with it that I will develop and respond to in the next chapter.

Possible worlds are maximal properties that a universe might have or, equivalently, maximal propositions. Each such proposition

is maximal in the sense that for every (actual) proposition, either it or its contradictory is entailed by it. But a proposition might be maximal in this sense while failing to be fully specific, where a proposition is fully specific only if for every existential proposition that it entails, it also entails a singular proposition that is a witness to that existential proposition.

Just saying this is a step toward reconciling simple modal claims about merely possible things with actualism. We can give truth conditions for statements such as "Saul Kripke might have had seven sons" without committing ourselves to the existence of anything that might have been one of Saul Kripke's seven sons. The statement is true if and only if there is a maximal proposition that entails the existential proposition that Saul Kripke had seven sons. This gives truth conditions for the possibility statement as a function of the inner proposition that is said to be possible, but we need our recursive semantics also to give the conditions under which the clause that expresses this inner proposition would be true *relative to a nonactual possible world*. It seems, however, that "Saul Kripke had seven sons" can be true, relative to a given possible world, only if seven singular propositions of the form "x is Saul Kripke's son" are also true with respect to that world, which is to say: only if seven singular propositions of this form are entailed by the maximal proposition that is that possible state of the world. But if maximal propositions can fail to be maximally specific, this condition will not be met.

Alan McMichael made the problem clear and precise in a classic criticism of actualist possible-worlds semantics¹⁸ by focusing on the problem of iterated modal propositions—for example, that Saul Kripke might have had seven sons, the last of whom was a plumber who might instead have become a lawyer. McMichael proved, using premises that he argued the actualist should accept, that an

¹⁸ McMichael 1983.

iterated modal claim such as this couldn't be true (assuming that no actual thing could have been Saul Kripke's seventh son). But as McMichael emphasized, one of his premises was that the abstract objects we are calling propositions exist necessarily, a premise the account I am promoting rejects. In response to this way out of the problem, McMichael argues that "to acknowledge [that the possible worlds which exist from the point of view of one world are distinct from those that exist from the point of view of another] is to give up the extensionality of possible worlds semantics. . . . But if we have to give up the extensionality of the possible worlds approach, we might as well do without it."¹⁹

McMichael also considers a response to the problem that gives a nonrealistic interpretation of possible-worlds semantics—one that rejects "the idea of there *really being* nonactual possibles" but employs "a semantics which includes *so-called* nonactual possibles," and he raises some problems for nonrealistic semantics. The problems he raises for reconciling actualism with possible-worlds semantics are serious and on target. I think they can be overcome, but doing so will require that I be more explicit about the way I want to use and interpret possible-worlds semantics. I will try to do this in the next chapter.

¹⁹ Ibid., 55.

{ 2 }

Merely Possible Possible Worlds

E. J. Lowe, in a general discussion of ontology, makes the following remark, in passing:

Many abstract objects—such as numbers, propositions and some sets—appear to be necessary beings in the sense that they exist “in every possible world.” . . . Indeed, possible worlds themselves, conceived of as abstracta—for instance as maximal consistent sets of propositions—surely exist “in every possible world.”¹

I suggested in chapter 1 that this is false. Possible worlds, in the sense in which it is reasonable to believe that there are many of them—the sense in which they are “conceived of as abstracta”—are contingent objects. It does not take a very sophisticated argument to make at least a *prima facie* case for this claim. It seems plausible to assume, first, that there are some propositions—singular propositions—that are object-dependent in the sense that the proposition would not exist if the individual did not. It also seems plausible to assume that there are some objects that exist only contingently and that there are singular propositions about those objects. These assumptions obviously imply that there are propositions that exist only contingently, and if possible worlds are maximal consistent propositions, or maximal consistent sets of propositions, it implies

Parts of this chapter were previously published as Stalnaker 2009. Thanks to Oxford University Press for permission to include them here.

¹ Lowe 1999, 248.

that there are possible worlds (or possible world-states) that exist only contingently. But despite the apparently compelling argument for this thesis, there are reasons to resist it, and it has been resisted. The thesis has some surprising and counterintuitive consequences, and there are some intuitively compelling arguments on the other side that we will have to consider. And if we accept the thesis, we need to consider its effect on a semantics for modal notions. My main aim in this chapter is to argue that we can reconcile the contingent existence of propositions with orthodox possible-worlds semantics, though the way of doing so that I will propose makes some concessions to the points made by Alan McMichael that I discussed briefly at the end of chapter 1. Our overall theory will, in a sense, involve a retreat, both from extensionality and from realism about possible states of the world, but it is a tactical retreat that, I will argue, preserves the virtues of both realism and extensionality.

Here is my plan for this chapter: I will first sketch a minimal theory of propositions—one that ascribes to propositions just the structure that anyone who is willing to talk of propositions at all must ascribe to them. In section 2 I will extend the minimal theory by adding some assumptions about the modal properties of propositions and possibilities, and in section 3 I will sketch a general model of logical space that makes room for merely possible possibilities. Sections 4 and 5 focus on the relation between models and the reality that they purport to model and on the extent to which our theory of propositions and possibilities provides a realistic semantics. In section 6 I will respond to some arguments against the thesis that propositions may exist contingently.

1. A Minimal Theory of Propositions

Propositions are the contents of speech and thought, and they are the objects to which modal properties like necessity and possibility are ascribed. There are many theories about what the things are that play these roles, but all will agree about a certain structure

of relations between propositions: they may be compatible or incompatible with each other, one may entail another, two may be contraries or contradictories, or necessarily equivalent. They are also objects that are true or false. Many of the properties of and relations between propositions are interdefinable. A minimal theory of propositions can make do with just two primitive properties: a property of *consistency* applied to sets of propositions, and a property of *truth* applied to propositions. Before stating the postulates of the theory, I will define four additional properties in terms of consistency that will be useful for stating the postulates:

(D1) A set of propositions Γ is *maximal consistent* iff it is consistent, and for every proposition x , if $\Gamma \cup \{x\}$ is consistent, then $x \in \Gamma$.

(D2) Two sets of propositions Γ_1 and Γ_2 are *equivalent* iff for every set of propositions Δ , $\Gamma_1 \cup \Delta$ is consistent if and only if $\Gamma_2 \cup \Delta$ is consistent.

(D3) A set of propositions Γ *entails* a proposition x iff $\Gamma \cup \{x\}$ is equivalent to Γ .

(D4) Two propositions x and y are *contradictories* iff they meet the following two conditions: (a) $\{x, y\}$ is inconsistent, and (b) for every consistent set Γ , either $\Gamma \cup \{x\}$ is consistent or $\Gamma \cup \{y\}$ is consistent.

Individual propositions are said to be consistent or equivalent (respectively) when their unit sets are consistent or equivalent, and an individual proposition x entails a proposition y iff its unit set $\{x\}$ entails y .

The postulates of the minimal theory are as follows:²

(P1) Every subset of a consistent set is consistent.

²These postulates differ from those given for a minimal theory of propositions in Stalnaker 1976. Earlier changes were in response to problems pointed out by Philip Bricker. Later changes were in response to problems pointed out to me by Damien Rochford.

- (P2) The set of all true propositions is maximal consistent.
- (P3) Every proposition has a contradictory.
- (P4) For every set of propositions Γ , there is a proposition x such that Γ is equivalent to $\{x\}$.
- (P5) Every consistent set of propositions is a subset of a maximal consistent set.
- (P6) Equivalent propositions are identical.

I take it that (P1) is unproblematic and needs no comment. (P2) is equivalent to the assumption that every proposition is either true or false. There may be applications of the notion of proposition for which this postulate might be denied (for example, a relativist semantics or a noncognitivist theory in which there are possibilities distinguished in thought such that there is no absolute fact of the matter which of them is actual), but I am going to ignore this complication. (P3) and (P4) are closure conditions on propositions. If propositions are something like truth conditions, then (P3) is just the assumption that if there is a certain truth condition, then there is also a condition that that condition *not* be satisfied, and (P4) is the assumption that for any set of different truth conditions, there is the condition that *all* of them be satisfied. It is not assumed that we necessarily have the resources to express all of these conditions—just that they exist.

(P5) might be denied on the grounds that propositions might be “gunky”: one might think that for every proposition x , no matter how specific, there are always further propositions that are incompatible with each other, but each is compatible with x . But while I don’t want to exclude the possibility that there are domains of propositions for which (P5) fails, I will restrict attention to domains of propositions for which it holds. I suggested in chapter 1 that one should think of possible worlds as cells of a partition of logical space rather than as points in the space—partition cells that

are fine-grained enough to settle all issues at hand. Equivalently, one can think of the domain of propositions as all of the propositions that concern the relevant subject matter. On this more deflationary and less metaphysical conception, the closure conditions (postulates (P3) and (P4)) are still reasonable assumptions as constraints on the set of propositions that concern a subject matter. And the problems concerning contingently existing propositions that are our concern will still arise.

Many accounts of what propositions are (for example, a Russellian theory of propositions or a Fregean account of Thoughts) would reject our last postulate, (P6), allowing for distinct propositions that are necessarily equivalent. But this postulate still belongs to a theory of propositions that is appropriately called "minimal" for the following reason: however propositions are individuated, all who are willing to talk of propositions at all should agree that propositions, as they understand them, *have* truth conditions and so that the equivalence relation in terms of which (P6) is stated is well defined. Our theory of coarse-grained propositions is minimal in the sense that it characterizes an entity that all theorists of propositions can agree about, even if they want to allow, in various different ways, for more fine-grained objects that determine propositions in this coarse-grained sense. For purposes of modal semantics, we need a notion of proposition that is the right grain to ensure that our semantics is compositional, and a notion of proposition satisfying (P6) seems to meet this condition, at least if we are not trying to represent intentional mental states, even if further distinctions may be needed for other purposes.

We can represent the minimal theory of propositions pictorially with a representation of logical space. Propositions are ways of dividing the space—they can be represented by subspaces, with compatible propositions represented by overlapping subspaces, and entailment represented by inclusion. The totality of propositions determines a maximally fine partition of logical space, with

the partition cells corresponding to maximal consistent propositions. Every proposition will correspond to a set of these maximal consistent propositions, and every such set will determine a unique proposition. This one-to-one correspondence was the ground for the identification of propositions with sets of possible worlds, or world-states, with world-states identified with maximal propositions, but this identification becomes problematic when we recognize that propositions themselves may exist only contingently.³ If we allow for this possibility, we must confront the question (to use familiar, perhaps metaphorical, jargon) of the identification of propositions across possible worlds, as well as some methodological questions about what we are doing when we use modal concepts to talk about the framework that we want to use to analyze and represent modal concepts.

2. The Modal Properties of Propositions

If we allow for contingently existing propositions and for the possibility of propositions that do not in fact exist, then there may be propositions that are *maximal consistent* in the sense defined but that are only contingently maximal and thus have the *potential* to be further refined. In chapter 1, I used a version of Kripke's dice example to give a toy illustration of this: two indiscernible dice are thrown; one lands 6, the other 5. Had that happened, there would

³ Aviv Hoffmann has provided a decisive refutation of the thesis that possible worlds can be identified with sets of maximal propositions, on the assumption that there are object-dependent singular propositions about contingently existing individuals. The contingency would proliferate, implying that purely existential propositions, in fact all propositions, were object-dependent. In Stalnaker 1976, I compare the two orders of analysis and note that given certain independently plausible assumptions, a minimal theory of propositions will yield the conclusion that there is a one-to-one correspondence between propositions and sets of maximal consistent propositions. But the move from this correspondence to an identification of the propositions with the corresponding sets ignores the modal properties of propositions.

have been a different possibility in which the one that landed 6 landed 5, and the one that landed 5 landed 6. Given that the dice are merely possible, no propositions that exist in the actual world can distinguish these two possible possibilities. There is, therefore, a maximal possibility that, had it been realized, would not have been maximal. Or consider a maximal proposition that entails that Saul Kripke had seven sons and that his seventh son was a lawyer. If this proposition were true, there would then exist singular propositions about that seventh son, including the proposition that he was a lawyer and the proposition that he was not a lawyer but a plumber. The first of these merely possible propositions (that he was a lawyer) would be true in w , even though it is not *actually* entailed by w , since it—the proposition—does not actually exist. The second of these singular propositions (that he was not a lawyer but a plumber) would be merely possibly true. There would also be a singular proposition about this seventh son that he did not exist—a proposition that would be false in the maximal counterfactual situation we are considering but (it seems reasonable to believe) merely contingently false. It would therefore be true with respect to certain possible worlds that would be counterfactual worlds if Kripke had had a seventh son. Among the possible worlds in which that singular proposition would be true is our actual world, at least assuming that in fact Kripke does not have seven sons. So there could have been a proposition that does not in fact exist but that if it had existed would have been true *with respect to the actual world*.

I used the term “possible possibilities” above, which is loose talk. There are no merely possible possibilities, of course, just as there are no merely possible people. What does exist is the possibility that there are possibilities, and propositions, that do not in fact exist, and the existence of these possibilities implies the existence of general propositions about propositions, for example, the (false) proposition that there exist singular propositions that witness to a particular (false) existential proposition. (These correspond to the

higher-order properties of containment properties that I discussed in chapter 1.) The problem is how to model the complex structure that relates propositions, including propositions about propositions, to each other.

Our minimal theory of propositions made no claims about the modal status of propositions—about their essential and accidental properties or about the relation between the propositions there are and the possibility of there being others. We might add two modal assumptions to the minimal theory:

(P7) There exists a proposition that *necessarily* entails all propositions.

(P8) For any set of propositions Γ , if Γ is consistent, then *necessarily* if Γ exists, then Γ is consistent.

The first of these modal postulates is guided by the idea that even though different possible situations may provide different resources with which to partition logical space, it is the same logical space that we partition, whatever possibility is realized. Our minimal theory already implied that there exists a (unique) proposition that is entailed by all propositions. The additional modal principle (P7) adds that this proposition *necessarily* has the property of entailing all propositions (and so necessarily exists). The second modal principle implies that all the basic propositional properties and relations (entailment, consistency, incompatibility, etc.) are essential properties and relations: for example, it follows from (P8) that for any propositions x and y , if x entails y then it is necessary that if x and y exist, then x entails y .

These additional principles may seem plausible enough, but is it legitimate for us to help ourselves to the notion of necessity in stating principles of the theory of propositions that will be the basis of our semantics for a modal language? I have disclaimed any pretension to be reducing modal to nonmodal concepts—I even claimed that it is a virtue of a theory that it avoids such a reduction. But I

also promised a vindication of possible-worlds semantics, a semantic theory that is presumed to get its explanatory power in part from its extensionality. As McMichael argued, “if we have to give up the extensionality of the possible-worlds approach, we might as well do without it.” But I hope to provide an interpretation of the orthodox semantics that retains the virtues of extensionality while also making use of primitive modality in the theory of propositions.

Kit Fine, following Arthur Prior, defended a thesis that he called “modalism”: “The ordinary modal idioms (necessarily, possibly) are primitive.”⁴ Fine takes modalism to be incompatible with the possible-worlds “analysis” of modal concepts, while acknowledging that possible-worlds semantics may have its uses if it is not regarded as providing an analysis. If by “analysis” one means an eliminative reduction, then I think most possible-worlds theorists (David Lewis aside) will agree with modalism, but one may still hold that possible-worlds semantics provides a genuine explanation, in some sense, of the meanings of modal expressions. The problem is to clarify the sense in which an explanation, short of reduction, is still an explanation. Consider the analogy with the semantics for first-order quantification theory: there are quantifiers in the semantical metalanguage, and of course the semantic “analysis” of the quantifiers provides no reduction of the concepts of existence and universality to anything more basic. But the theory still gives us a compositional semantics that sharpens and clarifies the structure of quantificational discourse—the ways that quantifiers interact with each other, as well as with names, predicates, and other logical operators. The project of reconciling possible-worlds semantics with intuitively plausible metaphysical commitments faces some special problems not present in the case of extensional quantification theory, but the fact that we need to use modal concepts to explain our primitives is not itself a problem.

⁴ Fine 2005, 133.

3. The Model of Logical Space

I will have more to say about primitive modality later in this chapter, but let me first elaborate a bit on our picture of logical space, partitioned by the maximal propositions: the way the space is partitioned is a contingent matter, depending on the resources available in the actual world—the world in which our theory of propositions is being stated. If one of the maximal propositions that is in fact false were to have been realized, there would have been different resources available, and logical space (the same logical space) would have been partitioned in a different way—perhaps more finely in some respects and more coarsely in others. (And when I say that different resources would have been *available*, I mean not just that those who are partitioning logical space would have had access to different resources; I mean that different resources would have existed.) We can model this conception of logical space with a set of points and a function taking each point to an equivalence relation that provides the maximal partition that *would* divide logical space, were that point to represent the possible state of the world that is realized. Now we know what the cells of our basic partition (for the actual world) represent: maximal consistent propositions. But what do the *points* represent? Specifically, what is the difference in what is represented by two distinct points within the same partition cell? The answer is that they all have exactly the same representational significance, but we need many of them in order to represent the way in which maximal propositions have the potential to be further refined. This potential is reflected in higher-order propositions about propositions and in the iterated modal propositions that were the basis of McMichael's challenge to what he called "atomistic actualism." Intuitively, one may think of the points as representations of possibilities, one of which *would* be maximal if the partition cell of which they are members had been realized. Which one of them would have been realized? There is no fact of

the matter about that, since each point within an equivalence class has exactly the same representational significance (in the actual world) as every other point within its equivalence class. But we need more than one such point in order to represent the different possibilities that *would* exist if that possibility were realized. To return to our toy example of the dice, we want to represent the fact that if our dice had existed, there would have been *two* possibilities in which one die landed 6 and the other 5, since had there been two such dice that landed 6 and 5, there would have been a distinct possibility in which they landed the other way. But since they don't in fact exist and are characterized generically as indiscernible, there is, in the actual world, only one maximal property (one 6, the other 5) to represent these two distinct possible possibilities.

Since a maximal consistent proposition will, by definition, decide *all* (actual) propositions, including the higher-order ones about what kinds of propositions would exist if that maximal proposition were realized, our framework must impose structural constraints on the family of equivalence relations to ensure that each of the points within any maximal equivalence class will decide *all* the higher-order propositions in exactly the same way. The technical details of these structural constraints are spelled out in appendix A.

4. Models and the Reality They Model

When we come to doing our modal semantics, it will be the points, not the maximal equivalence classes of points, that play the role of possible worlds. But now the attentive reader may be inclined to cry foul, complaining that she has been subjected to a bait and switch. I started by asking what possible worlds are and answered that they are properties that a universe might have—maximal properties or, equivalently (I proposed), maximal consistent propositions. I also promised a vindication of orthodox possible-worlds semantics, but now I am saying that the entities that are the “possible

worlds" in the models of our orthodox semantics are different from the maximal propositions that I began by identifying with possible worlds. While we might be persuaded to be realists about maximal propositions (the critic complains), it is another matter to be realists about these points that our models must use to make sense of iterated modal claims. Have we simply replaced the unacceptable nonactual possibles with *so-called* nonactual possibles (to echo McMichael), things such as numbers or sets that are *not* nonactual possibles but that are suitable to model them?

This complaint is fair enough, but I did warn you that there would be some concessions to the critics of actualist possible-worlds semantics and a tactical retreat from a fully realistic interpretation of that framework. And while the interpretation of the semantics that I want to give is perhaps not straightforwardly realistic, I will argue that it can answer the criticisms that McMichael makes.

McMichael had two reasons for finding a nonrealist strategy unsatisfactory. First, for any semantics that "contains nonrealistic elements, the problem will arise of distinguishing what aspects of the semantics are of genuine significance and what aspects are purely artificial. We will want a method for 'factoring out' the artificial aspects. But a nonrealistic semantics coupled with a method of 'factoring out' is just a realistic semantics."⁵ I think this is exactly right, but McMichael's own discussion points to the strategy for answering his challenge. He discusses the analogy with a relational theory of space. According to the relationist, there are really no such things as spatial locations—there are just spatial relations between things. But the best way to model the structure of spatial relations is with a (mathematical) space, made up of points (*so-called* spatial locations). One "factors out" the artifacts of the model—separates them from the realistic claims of the theory—by adding to the theory an equivalence relation between spatial models. Equivalent models

⁵ McMichael 1983, 63.

are those that differ in artificial ways but that agree in the realistic claims they make about the spatial relations between things. There are different versions of this kind of relational theory, made precise with different equivalence relations. One might say that all and only permutations of spatial points that preserve distance relations between points are equivalent representations.⁶ So a possible world in which everything, throughout history, is three feet to the north, or rotated forty-five degrees on a certain axis, from the way things are in the actual world is just a conventional redescription of the actual world. Or one might say that only *ratios* of distances need to be preserved. Adolph Grünbaum once discussed the verifiability and intelligibility of what he called the “universal nocturnal expansion hypothesis” that everything should instantaneously double in size. (Why this has to happen at night, I am not sure.)⁷ The question is one of the identification of spatial location, and spatial properties and relations, across possible worlds. Could I fix the reference of a spatial location (for example, the place of the center of mass of the sun, at a specified time) and then stipulate that a certain counterfactual situation should be one where some specified kind of event takes place *there*? The reference-fixing act presupposes that there exist locations reference to which can be fixed.

There are familiar questions of the identification of locations across time, as well as across worlds (or, equivalently, about the relations between spatial and temporal structures). Since the basic laws of Newtonian physics are invariant in all inertial frames, one might accept Newtonian physics while being a Galilean relativist, holding that there is no such thing as absolute velocity, only velocity relative to an inertial frame. An equivalence relation, defined by a class of permutation relations on spatio-temporal points, makes

⁶For simplicity of illustration, I am assuming the independence of spatial and temporal dimensions.

⁷See Grünbaum 1964. I believe that *Playboy* magazine once took note of this hypothesis, wondering (based just on its name, in the title of an article) what the hypothesis might be and what these philosophers were up to.

the notion of an inertial frame, and this relativistic thesis, precise. The Galilean relativist is, in a sense, anti-realist about spatial locations, but he can use the same basic framework for representing his theory as the Newtonian absolutist, and he can, as McMichael puts it, “factor out” the conventional, or anti-realist, aspects of his physics from the part about which he claims there is a fact of the matter. If we can be as precise as the Galilean relativist in stating the relevant equivalence relation, I think we will have given an adequate interpretation of our modal semantics.

I should add that there is more than just an analogy here. One may think of abstract spaces in general as representational devices for modeling properties and relations.⁸ The application to *physical* space and time is just the most prominent application. We will have more to say about space and spatial relations in the next chapter.

Another analogy that is relevant is utility theory: real numbers are used to specify utility values, which are intended to reflect in a systematic way the motivational states that dispose a rational agent to act in a range of different possible circumstances. But the numbers themselves are conventional: any positive linear transformation of a utility function is a representation of exactly the same motivational state. Even with fundamental measurement—quantities such as length and mass—there is a conventional element in the use of real numbers to represent the quantities, but the relevant equivalence relation is stricter, requiring the preservation of ratios, and not just ratios of intervals. In general, if one is trying to model a purely relational structure, the strategy of “factoring out” the artifacts of the model with an equivalence relation is a familiar one. According to an actualist, the facts about iterated modalities, at least those involving the possibility of things that do not in fact exist, are the kind of purely relational facts that should be modeled in this way.

⁸ See Hawthorne and Sider 2002 for an interesting discussion of abstract spaces used to represent properties and relations.

As McMichael says about the spatial analogy, “the problem of distinguishing . . . *invariant* features from artificial ones is just as important as finding a coordinate system that ‘works.’”⁹ The family of equivalence relations that is part of the modal theory of propositions I have sketched (and that is spelled out in more detail in the appendix) aims to make this distinction precise. If McMichael is right that “a nonrealistic semantics coupled with a method of ‘factoring out’ is just a realistic semantics,” then our interpretation of the orthodox possible-worlds semantics should count as a realistic one.

But is McMichael right about this? Is this indirect way of giving a semantics for a modal theory sufficient? The critic might say: “What I want to know is, what is there *really*, according to your theory?” The answer is: there are individuals—actual ones only—though what individuals there are is a substantive question, mostly empirical, and our abstract theory remains neutral about most of that. There are also properties, propositions, and relations (again, actual ones only). About them we have more to say, though we still remain largely neutral about what properties and relations there are (for example, about whether there are absolute spatial locations, or whether there are irreducible mental properties). But there are (according to our theory) not only properties of individuals¹⁰ but also higher-order properties and relations: properties of properties, relations between properties, and relations between properties and other things. I talked about some of these higher-order properties and relations in the discussion of containment properties in chapter 1: there was, for example, a correspondence relation between generic and specific properties. There will be a correspondence relation like this between existential propositions and their witnesses. Salient among the higher-order properties and relations are

⁹ McMichael 1983, 63.

¹⁰ By “individual” here, I mean things that are not themselves properties, propositions, or relations.

propositional functions. A function, in general, is a special case of a relation, so a propositional function is a kind of relation between an individual and a proposition. And there are also functions from individuals to propositional functions, functions from individuals to functions from individuals to propositional functions, and on up from there. (See appendix B for a digression on propositional functions and a way to do the modal semantics that does not involve quantification, in the metalanguage, over merely possible individuals.)

So there are properties and relations. We can say, in the context of Kripke models, how properties, relations, and propositional functions are modeled: properties, for example, are modeled as functions from possible worlds (points) to subsets of the domain of that world. But what *are* properties, really, in themselves? As I said in chapter 1, properties are to be understood in terms of the way the world would have to be for them to be instantiated. The concept of a property is a basic concept, not reducible to something else, but it is not an isolated concept: we can say a lot, not just about particular properties (by elaborating, say, about what a thing must be like to be a donkey, to be yellow, or to be soluble) but also about how properties and relations are related to each other—about what their role is in a complex relational structure. How do we do that? By using models, such as standard Kripke models, together with a family of equivalence relations on the points in the model that spell out what features of the models are artifacts and what parts are features of the reality being modeled.

In characterizing possible worlds-states as a kind of property, I emphasized that they are therefore not (or at least not essentially) representations. But of course we can theorize about properties, propositions, and possible states of the world only by representing them (just as an astronomer can theorize about planets and stars only by representing them and their properties). Unlike the possible states of the world—maximal ways a world might be—that

are being modeled in our semantic theory, the points (the “possible worlds”) in a Kripke model *are* representations—representations of the properties that I claim possible states of the world are. The whole Kripke model represents not just these properties but also a structure of relations between these properties (the possible states of the world) and between them and other things. The points themselves are not properties—they are points in an abstract space that are being used to represent possible states of the world. So I won’t even complain if you call the points in a Kripke model “ersatz possible worlds” or better, “ersatz possible *states* of the world,” since it is ways a world might be that they represent.

Aviv Hoffmann, in a commentary on the paper that this chapter draws on,¹¹ argues that I have not really defended the thesis that propositions and possible world-states are contingent entities, but rather the different thesis that the things I call propositions and possible worlds are only contingently propositions and possible worlds. According to Hoffmann, what I have shown, at best, is that being a proposition may be a contingent property but that on my account, the things themselves that are contingently propositions are still necessary existents. The argument is based on a principle of modal set theory, which should be uncontroversial: a set exists in a given possible world if and only if all of its members exist at that world. Hoffmann argues that since for each point in logical space it is necessary that there exists some proposition containing that point, it follows that all the points, and so all the subsets of the set of all points, necessarily exist. His conclusion is that while I can deny that certain subsets are propositions, relative to certain possible worlds, I cannot deny that they exist, relative to that world. But I think Hoffmann’s argument conflates the propositions and possible states of the world with the sets of points of a space that represent them. The points are elements of an abstract space.

¹¹ Hoffmann 2010.

Whether the space itself, with its points, is a necessary existent or not is an independent question. Even if mathematical spaces of the right structure existed only contingently, one could still use them to represent propositions that entail the nonexistence of the space itself. On the other hand, if the space and all of its points are necessary existents, that does not prevent us from using them to represent contingently existing things. (We cannot, for example, infer from the fact that Babe Ruth is represented by the number 3 that the Babe himself is therefore an eternal and necessary being.)

5. Nonrealistic Semantics

I have argued that our overall semantic theory is, in the sense that matters, a realistic one, while acknowledging that there is a sense in which we are not realists about the possible worlds, and possible individuals of the Kripke models used to represent the ways things might be. I have tried to meet, in a precise way, McMichael's challenge to "factor out" the nonrealistic elements of the theory, and if I have succeeded, then McMichael's more general concern about a nonrealist semantics should not apply to the account I am promoting. But let me consider that concern and distinguish more explicitly the account I want to defend from one that simply uses a domain of "so-called" possibilia as surrogates for what might exist but does not.

McMichael concedes that a nonrealistic semantics that uses some arbitrary surrogates for possibilia might be sufficient if our project is only to do semantics for modal *logic*. He suggests that a nonrealistic semantics might give an adequate account of the notions of validity and satisfiability that are appropriate for modal language, but "what we are ultimately interested in," he says, "is to give truth conditions for some nice modal fragment of a natural language," and a Kripke semantics, by itself, "does not supply conditions for *truth* [as contrasted with conditions for *validity*]."

If what we are interested in is *truth* conditions, he claims, then we need to answer questions about “the number and relationships of nonactual possible individuals.”¹²

I don’t think the distinction between natural language semantics and semantics for the formal language of modal logic is relevant to the issue McMichael is raising. In a robustly realistic theory, the languages in which we formulate our philosophical commitments might be regimented formal languages rather than fragments of natural language. The issue he is raising, I think, is about how to understand the models that are used to interpret the language. I take the point to be something like this: an *intended* model, in a robustly realist semantics, is not a *representation* of the subject matter of one’s theory; it is the subject matter itself. If I formulate a theory that I wish to defend—if I claim that my theory is true—then in the intended model for my theory, the domain will be the things I claim to be talking about and not substitutes for them. In the intended model for a theory about DNA molecules, for example, the domain (which is part of the model) contains DNA molecules themselves and not models (perhaps made out of Tinkertoys) of them.

Unintended models, as McMichael suggests, have their place. Suppose I have a friend who has an elaborate theory about elves, sprites, fairies, and leprechauns. I don’t accept his theory and in fact would have a hard time specifying, in a precise way, the *truth* conditions for his theory—saying exactly what the world would have to be like for his theory to be true. But if my task is just to assess the validity of his reasoning when he expounds his theory and draws conclusions from its basic principles, I will have an easier time. I might regiment his theory in a first-order (extensional) language and specify a model for it, perhaps using a set of natural numbers for the domain (to be the *so-called* elves, sprites, fairies, and

¹² McMichael 1983, 63.

leprechauns). My model is not my attempt to specify my friend's *intended* model, but if my interpretation makes the sentences he uses to state the theory come out true in the model, that might be good enough for my purpose, which is just to assess his reasoning. I will not, however, have provided a semantics for my friend's theory that is realist, in any sense.

I agree with McMichael on this general point but would argue that the relationship between a theory and its subject matter may be more complex than the simple story suggests. Models (in the model theorist's sense of the term) may be used in different ways in one's representation of the subject matter of a theory; in particular, models may be interposed between a language used to talk about some domain and the domain that the theory talks about. That, I want to suggest, is what is going on when one uses a model, plus an equivalence relation, to represent a structure of relations, as in a relational theory of space and time. A theorist who represents her theory of space and time in this way may be a realist about spatio-temporal relations and about the physical objects that exemplify such relations, even if she is not a realist about spatial locations themselves. As I have said, the merely possible individuals, and the points in logical space used in Kripke models as I am interpreting them, are like the spatial points in a relativist's model of spatial structure. The intended subject matter of our modal theory consists of the actual individuals, the (actual) properties and relation that they might exemplify, and the (actual) higher-order properties and relations that might be exemplified by properties, relations, and propositions, as well as by individuals. About all these things, our theory can be resolutely realistic.

But while I agree that to be a realist, in the sense I want to be a realist, is to accept a commitment to the existence of the things in the domain of the intended model of one's theory (allowing that models may also play an intermediate representational role), this is not necessarily to accept a commitment to the comprehensiveness

of one's theory—to accept (to echo the words of Barwise and Perry quoted in chapter 1) that there is a unique intended model of super-reality that will provide us with a complete inventory of all the things, properties, and relations there are in the universe. I am inclined to be skeptical about the positive answer to the contentious philosophical question whether it makes sense to quantify over absolutely everything, but this is a separate question on which the kind of modal theory I am defending remains neutral.

If you press me on the question, how many points there should be in a model that represents the metaphysical possibilities, including the possibilities of things that might exist but don't, I am inclined to answer, as many as you need to model the modal propositions that you want to model. The answer to the question, how many is that, will depend both on the expressive resources of the language one is using the models to interpret and on one's metaphysical views about what is possible. The abstract semantic theory won't answer those questions, but I hope at least that it will help clarify what is being asked.¹³

6. Objections and Replies

I will conclude this chapter by looking at some of the problematic consequences of the thesis that propositions may exist merely contingently, but first let me remind you of the *prima facie* case for this thesis.

It is motivated by just two simple assumptions. The first is a doctrine Alvin Plantinga calls *Existentialism*: “a singular proposition is ontologically dependent on the individuals it is directly about.”¹⁴ The second is the claim that there are some things that exist only

¹³ Thanks to an anonymous reader for questions and comments that helped me clarify my response to McMichael's worries about anti-realist interpretations of the modal semantic.

¹⁴ Plantinga 1983, 160.

contingently. The second of these assumptions seems to require little explanation or defense, although some philosophers have denied it, as we will see. It seems at least *prima facie* reasonable to take it to be a Moorean fact that people and ordinary physical objects are things that might not have existed. But what about the first assumption? There are different accounts of what propositions are that might motivate this thesis. If you think of a singular proposition as a kind of Russellian proposition, an ordered sequence containing the individual, along with properties and relations, as constituents, then it is natural to think that the existentialist thesis must be true, since it is natural to believe that sets and sequences are ontologically dependent on their elements. But even if one is presupposing, as I am, a coarse-grained conception according to which propositions are individuated by their truth conditions, it seems *prima facie* plausible to think that propositions about particular individuals are ontologically dependent on the individuals they are about. On the coarse-grained conception, propositions are truth conditions, and the truth condition for a singular proposition is a condition that the world must meet (for the proposition to be true) that essentially involves the individual that the proposition is about. It seems reasonable to believe that a condition that depends for its satisfaction on the way Socrates is requires, for its existence, the existence of Socrates. While our propositions are not complexes with properties and individuals as constituents, we retain the idea that propositions are built out of the materials we find in the actual world. Any actualist must accept this, but what materials one thinks there are will depend on one's metaphysical and empirical beliefs. One can reconcile actualism with a rejection of object-dependence, if one is willing to make certain metaphysical commitments. One way to do this is to hold that there are qualitative conditions that are necessary and sufficient for the existence of a particular individual, so that singular propositions about actual or possible individuals are reducible to purely qualitative propositions. We will

discuss this strategy in chapter 3. Alvin Plantinga adopts a different strategy, which is to hold that for each individual, there is an individual essence, or haecceity, that exists independently of an individual that exemplifies that individual essence. So for Plantinga, while there may be no actually existing thing that would be Saul Kripke's seventh son if he had seven sons, there do exist properties, probably infinitely many of them, that would suffice to individuate each of the possible individuals who, in each of the possible worlds in which Kripke had seven sons, would have been his seventh son. One might use an identity property to fix the reference of a term referring to an individual essence, for example, the property of being identical to Obama. But the assumption is that while reference may be fixed in this way, the individual essence itself is a property that exists independently of the object used to fix the reference. This may seem an extravagant metaphysical commitment, but Plantinga has an argument against the thesis of object-dependence, or existentialism, which commits him to it.

Plantinga's argument has five premises:

- P1. Possibly, Socrates does not exist.
- P2. If P1, then the proposition *Socrates does not exist* is possible.
- P3. If the proposition *Socrates does not exist* is possible, then the proposition *Socrates does not exist* is possibly true
- P4. Necessarily, if *Socrates does not exist* had been true, then *Socrates does not exist* would have existed.
- P5. Necessarily, if *Socrates does not exist* had been true, then Socrates would not have existed.¹⁵

The conclusion drawn from these premises is as follows:

- C. It is possible that both Socrates does not exist and the proposition *Socrates does not exist* exists.

¹⁵ Plantinga 1983.

The first three premises obviously entail that *Socrates does not exist* is possibly true, and this, together with P4 and P5, entails C. The latter inference has the form:

$$\Diamond P, \Box(P \rightarrow Q), \Box(P \rightarrow R), \text{ therefore, } \Diamond(Q \& R).$$

This is an inference that is valid in any normal modal logic, so the argument as a whole is valid.

Plantinga takes premise P1 to be uncontroversial, not anticipating Williamson's reason for rejecting his conclusion, and everyone will accept P5. But Plantinga notes that each of the other premises has been denied (by Larry Powers, Arthur Prior, and John Pollock, respectively), and he considers three different defenses of object-dependence that choose one of these premises to reject. This way of setting up the problem exaggerates the differences between the three responses to the argument that he considers, since (I will argue) there is an equivocation in the consequent of P2 (and the antecedent of P3), and the choice of which premise to reject depends on how that equivocation is resolved. The responses that Plantinga calls "Priorian existentialism," which rejects P2 and "Powersian existentialism," which rejects P3, are different only in that they resolve the equivocation in different ways. The third response that Plantinga considers, "Pollockian existentialism," which rejects P4, also turns on the distinction between the two ways of understanding truth, though since I don't think one can get to the second stage of the argument in any case, the rejection of P4 on one (less natural) interpretation is not necessary to defeat the argument.

To bring out the different interpretations of the clause that is the consequent of P2 and the antecedent of P3, let me introduce some notation. First, I will use ' π ' as a term-forming operator on sentences. For any sentence ϕ , ' $\pi\phi$ ' will denote the proposition expressed by ϕ . So if S abbreviates the sentence "Socrates does not exist," then ' πS ' will denote the proposition that Socrates does not exist. Second, I will use the letter 'T' to be the monadic truth

predicate, applied to propositions. Third, I will use a binary predicate, 'T' relating propositions, for "entails" (as defined in our minimal theory of propositions). So 'Ixy' says that proposition *x* entails proposition *y*. Fourth, I will use a variable-binding abstraction operator to form complex predicates. ' $\hat{x}(Fx \vee Gx)$ ' will be a monadic predicate that will have, as its extension, the individuals that are in the extension of either *F* or *G*. For present purposes, the variables '*x*' and '*y*' will range over propositions generally, but the variable '*w*' will be restricted to possible worlds, understood as maximal propositions.

Plantinga's premises P2 and P3 are stated in terms of a predicate of propositions, "is possible," which might be defined in terms of truth or entailment in several different ways. Here are two definitions, which may not be equivalent.

$$\begin{aligned}\text{'possible}_1' &=_{\text{df}} \hat{x} \Diamond Tx \\ \text{'possible}_2' &=_{\text{df}} \hat{x} (\exists w) Iwx.\end{aligned}$$

The first predicate applies to propositions that are possibly true *in* some possible world, while the second applies to propositions that are true *of* or *entailed by* some possible world. The two definitions will be equivalent if all propositions exist necessarily, but not if some do not.

If one understands the predicate of possibility, as it occurs in the premises of Plantinga's argument in the first way, then the defender of object-dependence should reject P2 but accept P3 (opting for the Powers response). On the other hand, if one understands the predicate in the second way, then the defender of object-dependence should accept P2 but reject P3 (opting for the Prior response). The singular proposition *Socrates does not exist* is a proposition that will be true *of* or *entailed by* only possible worlds in which that proposition does not exist. Since we are agreeing with Plantinga that nothing can be truly predicated of something that does not exist, the truth predicate will not apply to any proposition in a possible world in which that proposition does not exist.

The Pollock response rejects P4. The most natural way to take P4 is to take the antecedent of the conditional at face value as an application of the predicate of truth to the proposition. But there may be some uses of the word "true" that should be understood as treating it like a redundant operator, its role being either rhetorical or to help mark a scope distinction. But given that there is a way to understand P5 so that it is true, it does not really help to find an interpretation according to which it is false. So I think it is the diagnosis in terms of the equivocation in P2 and P3 that shows where the argument fails.

I have argued that Plantinga's argument can be resisted, even by those who accept its premises, but the thesis that some propositions and possible states of the world exist only contingently does have some discomfiting consequences. To start, I want to note that any account of propositions and possible worlds, the one I am defending or any alternative to it, that allows for contingently existing propositions will require the distinction that I appealed to in my discussion of Plantinga's argument between what Kit Fine has called inner and outer truth: there will be propositions that are true *of* or *at* or *with respect to* a possible world, while not being true *in* that possible world. For a proposition to be true *in* a possible world is for it to have, in that world, the property of truth. For a proposition to be true *of* a possible world is for it to stand (in the actual world) in a certain relation (the entailment relation) to that possible world. That these two notions come apart follows from the following assumptions: (1) some propositions exist only contingently; (2) every proposition has a contradictory (this is a postulate of our minimal theory of propositions); and (3) necessarily, only existing things have properties, and in particular, only existing propositions have the property of truth. Here is the argument. By (1), there is a proposition *x* that exists only contingently, which means that there will be a possible world-state *w* that does not include *x* in its domain. But then by (2), its contradictory will also not be in the domain of *w*, and so by (3), neither *x* nor its contradictory

will have, in w , the property of truth. But world-states are, by definition, maximal, and so for any proposition, w will entail either the proposition or its contradictory. So since w will entail either x or the contradictory of x , there will be a proposition that is true *of* that world-state but not true *in* it.

But if there are cases of propositions that are true *of* or *at* a certain possible situation but not true *in* that situation, because they do not exist there, there will be violations of the necessitation of a simple truth schema—the schema for sentences of the form “ ϕ if and only if it is true that ϕ .” In the notation we have been using, the schema and its necessitation are $\phi \leftrightarrow T\pi\phi$ and $\Box(\phi \leftrightarrow T\pi\phi)$. In possible worlds *of* which ϕ is true but in which the proposition does not exist, $T\pi\phi$ will be false in virtue of the reference failure, in that possible world, of the term $\pi\phi$. Furthermore, there will be a divergence between the *entailment* relation and necessary truth preservation; that is, there will be counterexamples to $(Ixy \leftrightarrow \Box(Tx \rightarrow Ty))$. Proposition x might entail proposition y , even if y might not exist in a possible world in which x is true, in which case the equivalence will fail in the left-to-right direction. (An example: The proposition that no one is immortal entails the proposition that it is not the case that Obama is immortal. But if Obama had not existed, the proposition that he was [or that he was not] immortal would not exist, and so the proposition that it is not the case that he is immortal would not be true. But it might still be true, in such a counterfactual situation, that no one was immortal.)

These consequences might lead one to be skeptical of the true-in/true-at distinction. Timothy Williamson rejects it, as applied to propositions, though he acknowledges that it makes sense for sentences and other objects that express propositions.¹⁶ His claim is that the only way to understand “true-at” is as “true-in.” Since he also accepts object-dependence, or existentialism, he draws

¹⁶ See Williamson 2001.

the conclusion that everything exists necessarily and that nothing could exist except what does exist. I find his reasons for rejecting the distinction unpersuasive, but one of his worries points to the general issue we've met before about the extent to which the concepts of our theory are explanatory. He puts the problem as a worry about circularity. Suppose we define a possible world as a consistent and complete set of propositions, where a set Γ of propositions is consistent if and only if for any proposition y , if there is a valid argument from Γ to y , then there is not a valid argument from Γ to the contradictory of y , and Γ is complete if and only if for any y , there is a valid argument either from Γ to y or from Γ to the contradictory of y . (If we understand "validity" as entailment, then this is exactly what we have done in our minimal theory of propositions, so no problem so far.) But now he proposes that we explain the notion of validity (entailment) as *necessary* truth preservation. Possible worlds are thus explained in terms of the notion of validity, which is explained in terms of truth and necessity. But necessity is explained in terms of truth at (or with respect to) all possible worlds.

In our minimal theory of propositions, we did use a necessity operator in the metalanguage, and even though using this resource is clearly incompatible with a use of possible worlds to provide a reductive analysis of possible worlds, I argued that it did not make our semantics unacceptably circular (though I offered no general account of what makes a nonreductive theory explanatory). But I did not use our metalinguistic necessity operator to analyze the notion of entailment in the way that Williamson suggests. I took *consistency* as a primitive of the theory of propositions, though we could have begun with entailment and defined consistency in terms of it. A more austere theory of propositions might have just one primitive—a truth predicate—defining entailment (in the modal metalanguage) as follows: proposition x entails proposition y if and only if necessarily, if x is true, y is true. But as I have

noted, our account rejects this analysis, not because it makes the explanatory circle too tight (though this might be reason enough) but because it gets entailment wrong. There is a dialectical standoff here, since Williamson is right that true-of collapses to true-in, if we accept the more austere theory of propositions. I acknowledge that it seems *prima facie* plausible to identify entailment between propositions with the necessity of the (material) conditional that if one is true, so is the other. But if we acknowledge the possibility of contingently existing propositions, it is easy to understand why and how this identification may fail. Given the metaphysical implausibility of the ontological commitments we must undertake to avoid contingently existing proposition, it seems to me that giving up this identification is a small price to pay.

I don't have an argument against the existence of the kind of individual essences for nonexistent things that Plantinga believes in, nor do I have an argument against the existence of a vast population of actual things that could have been living people and other material beings but actually are not and reside in some realm outside of space and time. The famous incredulous stare that David Lewis took to be the strongest argument against his modal realism is good enough for me, not only for Lewis's modal realism but also for the metaphysical commitments of Plantinga and Williamson, which seem to me to have about the same degree of *prima facie* plausibility. My aim has been to show that one can at least give a coherent account of modality that allows for an expansive view of what is possible—one that accords with our pre-theoretic modal beliefs—without committing oneself to an excessively extravagant view of what actually exists. The modal framework is supposed to be neutral, allowing for the kinds of ontologies that Williamson and Plantinga endorse, as well as for those that accord more closely with common opinion. But if we succeed in rebutting some unsound arguments in support of such ontologies, I think the temptation to believe in them should go away.

Still, if all we do, in our metaphysical argumentation, is rebut arguments for accepting certain metaphysical theses, there will remain a question about what it is that makes one or another of the available alternative metaphysical doctrines correct, or at least worthy of acceptance. I won't have a general answer to this question, which continues to puzzle me, but in the next chapter I will consider some of the arguments for and against some contrasting views about the relation between individuals and their properties.

What Is Haecceitism, and Is It True?

Alvin Plantinga's way of avoiding contingently existing propositions was to hypothesize a domain of primitive but necessarily existing individual essences—actual entities that correspond, one-to-one, with the merely possible entities that are hypothesized by a possibilist metaphysics. Timothy Williamson's way of avoiding them was to deny that there could have been anything other than what actually exists and to claim that nothing there is could have failed to exist. There is a third strategy that has many precedents in the tradition and that some may consider more plausible, though it also involves a controversial metaphysical commitment. This is the *anti-haecceitist strategy*, which holds that a purely qualitative characterization of a possible world would be a complete characterization; there is, on this view, nothing to being a particular individual other than being an individual meeting certain qualitative conditions. My aim in this chapter is to explore this strategy, first trying to sharpen the doctrine and its haecceitist alternative and then considering some of the arguments for and against the alternative theses.

Before I begin I should note that it is not obvious that even if we were to accept the anti-haecceitist metaphysics it would give us a way of avoiding commitment to contingently existing propositions. Singular propositions about contingent concrete particulars are the most salient *prima facie* examples of contingently existing propositions, but I see no reason to assume that even contingent general propositions, such as universal and existential generalizations

involving purely qualitative properties, might not be contingently existing propositions as well. Consider, for example, color properties—paradigms of purely qualitative properties. If there are metaphysically possible worlds with radically different physical laws, perhaps worlds without light or other kinds of electromagnetic radiation, and without conscious creatures with anything remotely like our visual experience, then there will presumably be worlds in which nothing is or could be colored. It would also be reasonable to conclude that the color properties themselves would not exist in such a world. Just as the truth conditions that we identify with singular propositions about Socrates are conditions on Socrates, so the truth conditions for statements about the colors of things seem to be conditions on the color properties—conditions in which *they* would be exemplified, or not. So if such properties are contingent existents, then even if singular propositions about people and physical objects were in some sense reducible to purely general propositions, we might still face the problem of contingently existing propositions and still need the distinctions that allow us to make sense of them. The structure of the framework I sketched in the last chapter (and in more detail in appendix A) is completely general; though I motivated the framework with cases of true existential propositions that have no true witnesses, it will apply to any case of contingently existing propositions, whatever the source of the contingency. But even if rejecting haecceitism would not avoid a commitment to contingently existing propositions, the issue of haecceitism is of interest in its own right, and discussions of its pros and cons will provide a good illustration of the dialectic of metaphysical argument. Though the question begins with a relatively abstract and narrowly focused issue, it quickly connects with a range of other metaphysical issues, concerning, for example, space, time, and counterfactual conditionals. And as is usually the case in metaphysical disputes, substantive issues are intertwined with issues about semantics and intentionality.

Here is the plan for this chapter: I will start with a preliminary attempt to characterize the contrasting theses, haecceitism and anti-haecceitism, discussing some of the different ways the contrast has been drawn in the literature. Then I will use the framework I sketched in the last chapter to try to sharpen the theses and to bring out their connection with some related metaphysical issues. The contrast between the two theses presupposes a distinction between properties and relations that are purely qualitative and those that are not, and I will argue that the content of the thesis of haecceitism, and the arguments for and against it, depend on the way this problematic distinction is understood. I will conclude with a discussion of the relationship between anti-haecceitism and the principle of the identity of indiscernibles, and of David Lewis's attempt to explain some of the haecceitist intuitions with a metaphysically cheap version of the doctrine.

1. David Kaplan's Version of Haecceitism

The jargon is of course of medieval origin, but the term 'haecceitism' was introduced into contemporary discussions of modal metaphysics by David Kaplan.¹ He characterized the doctrine in highly metaphorical terms, but the basic idea of his way of drawing the distinction was that the haecceitist allows, while the anti-haecceitist rejects, the irreducible identification of individuals across possible worlds. According to Kaplan's anti-haecceitist, individuals in different possible worlds may be "clothed in attributes which cause them to resemble one another closely. But there is no metaphysical reality of sameness or difference which underlies the clothes." The haecceitist, in contrast, holds that for individuals in different possible worlds, "a common 'thisness' may underlie extreme dissimilarity or distinct thisness may underlie great resemblance." Kaplan concludes his characterization of the contrasting doctrines with

¹ Kaplan 1975. The quotations in this paragraph are from pages 722–23.

this puzzling remark: "Although the Anti-Haecceitist may seem to assert that no possible individual exists in more than one possible world, that view is properly reserved for the Haecceitist who holds to an unusually rigid brand of metaphysical determinism." There may be a difference between Kaplan's anti-haecceitist and his rigid metaphysical determinist, but if there is it will require further explanation that will involve disentangling metaphysical and semantic issues.

Let me elaborate on this last point before turning to some different characterizations of the haecceitist doctrine. Consider David Lewis's metaphysical theory, which claims explicitly that "no possible individual exists in more than one possible world." In terms of Kaplan's characterization, should Lewis be regarded as a "haecceitist who holds to an unusually rigid brand of metaphysical determinism" or as an anti-haecceitist? Lewis's theory has been criticized as implying that everything has all of its properties essentially: that, for example, I could not have had something different for breakfast this morning than what I in fact had. But Lewis rejects this criticism, arguing that I have (in the actual world) modal properties that are defined in terms of the properties that my counterparts have in other possible worlds. So it is actually true that I might have had something different for breakfast this morning in virtue of the fact that one of my counterparts, in another possible world, had something different for breakfast on a morning that was a counterpart of this morning. The critic responds that this revisionary semantics just hides the implausible metaphysics. Lewis rejoins that his semantics gives the correct interpretation of our ordinary modal claims and succeeds in reconciling his metaphysics with common opinion about statements attributing modal properties. Now suppose a philosopher who accepted Lewis's thesis that "no possible individual exists in more than one possible world" were to reject the counterpart semantics, agreeing with the critic that the thesis implies that everything has all of its properties essentially, so it is right that I could not have had a different breakfast this morning.

This philosopher is defending “an unusually rigid brand of meta-physical determinism,” though she agrees with Lewis that there is a possible world very similar to this one in which someone very similar to me had a different breakfast on a day that is very similar to today, and we all can agree that to say this is to ascribe a (relational) property to me (the property of being an x such that there is a world in which someone very similar to x had a breakfast on a day similar to today that was different from the breakfast that x had today). What the rigid determinist rejects is the Lewisian counterpart semantics that represents this relational modal property of me in a certain way. Now it seems to me clear that the difference between Lewis and the haecceitist rigid determinist is a difference about how to do semantics and not a difference of metaphysics. Is the metaphysics of Kaplan’s anti-haecceitist different from Lewis’s? It does not seem so, since Kaplan’s anti-haecceitist holds that “there is no metaphysical reality of sameness or difference” of individuals in different possible worlds other than the sameness or difference in their qualitative properties.

There is a genuine issue about how to do the semantics for modal predication, and it may be that the assessment of Lewis’s semantics cannot be separated from the assessment of his metaphysics. But Kaplan’s anti-haecceitist and his rigid determinists seem to be in complete agreement about the metaphysics; the issues that remain for them concern how to describe the reality they agree about.

2. Robert Adams’s Version of Haecceitism

Robert Adams characterizes the haecceitist doctrine in a different way, without relying on the imagery of possible worlds.² He begins with the idea of a “thisness,” which is a monadic property that is defined in terms of the identity relation. For any individual—Osama

² Adams 1979.

bin Laden, for example—we can define the property of being identical to that individual. Osama, and only Osama, has the property of being identical to Osama. Now haecceitism (on Adams's characterization) is the thesis that thisnesses are not purely qualitative properties and are not reducible to purely qualitative properties and relations. The doctrine, defined this way, is of course only as clear as the distinction on which it rests. As Adams says about his defense of primitive thisness, “without the distinction between qualitative and nonqualitative, the subject of this paper does not exist.”³ I will consider his explanation of the qualitative/nonqualitative distinction below, but first I will sketch a third way of characterizing the thesis by a self-proclaimed defender of the anti-haecceitist doctrine, David Lewis.

3. David Lewis's Version of Haecceitism

Lewis characterizes anti-haecceitism as a supervenience thesis, so his first task is to circumscribe the supervenience base. There are two ways, Lewis says, that possible worlds might differ:

- (1) Worlds might differ in their qualitative character; or . . . in the qualitative character they ascribe to the concrete world. That is to say, they might exhibit or represent different patterns of instantiation of the natural intrinsic properties and external relations, and might thus disagree about just what kinds of things there are or about how things of various kinds are spatiotemporally arranged and causally related. . . . (2) Also, worlds might differ in what they represent *de re* concerning some individual: this-worldly individuals at least, and also other-worldly individuals if such there be. . . . There are some [possible worlds] . . . which represent *de re* of Humphrey (the

³ Ibid., 7.

Humphrey of our world), that he wins the presidency and there are others that represent, of him, that he loses.⁴

The anti-haecceitist doctrine, according to Lewis, is the thesis that the second kind of difference between possible worlds supervenes on the first.

Like Adams, Lewis relies on a distinction between properties and relations that are qualitative and those that are not. But what is distinctive about Lewis's characterization is the way he specifies what it is that (according to the anti-haecceitist) supervenes on the qualitative character of the world. For Lewis, the issue is not about the distinctions there are to be made between the possible worlds. On his account, haecceitistic differences, if there are such differences, are not differences between the possible worlds themselves but differences in the way possible worlds are used to represent individuals in other possible worlds. One might have expected a more straightforward supervenience thesis: that all differences between possible worlds supervene on purely qualitative differences. But Lewis explicitly rejects the identification of his thesis with this simpler supervenience claim: "Anti-haecceitism is neutral about whether there are qualitatively indiscernible worlds."⁵ I think it is Lewis's modal realism that accounts for this idiosyncratic characterization. Lewisian worlds are concrete particulars, with particular world-bound individuals as parts. Given this way of understanding what a possible world is, it is consistent to allow for distinct but absolutely indiscernible possible worlds while at the same time insisting that worlds that are indiscernible in all qualitative respects are therefore indiscernible absolutely. But it is not easy for Lewis to specify the kind of difference between worlds (according to his version of anti-haecceitism) that supervene on

⁴ Lewis 1986, 221.

⁵ *Ibid.*, 224.

the qualitative differences. Lewis's general metaphysical picture presupposes anti-haecceitism, as Kaplan defines it. For him, haecceitistic differences are differences in the way possible worlds are used to represent the properties of individuals in different worlds, not differences between the worlds themselves. But I think his way of characterizing the doctrine blurs the difference between metaphysical questions about what kinds of distinctions there are to be made between the possibilities and questions about how we can and do talk and think about the possibilities. In any case, I want to characterize haecceitism and its contrast within the actualist framework that takes possible worlds to be properties that a world might have. The idea of distinct but absolutely indiscernible Lewisian worlds may be intelligible, but there is no sense to be made of distinct ways the world might be such that being one of these ways is absolutely indistinguishable from being the other. So I am going to stay with the simpler supervenience thesis as a characterization of anti-haecceitism: it is the thesis that all differences between possible states of the world supervene on qualitative differences. But this leaves us with the question of how to make the distinction between the qualitative and the nonqualitative.

4. Distinguishing Suchness from Thisness

Both Adams and Lewis rely, to some extent, on a presumed intuitive grasp of the distinction between the qualitative and the nonqualitative (or to use Adams's language, between suchness and thisness), a grasp that I am not sure we have. Before looking at what they both say about the distinction, let me emphasize the diversity of the properties and relations that we may presume to be qualitative in the broad sense that is necessary to be a general contrast with properties and relations that involve particular individuals. First, the paradigm cases that most obviously deserve the name "quality" are

what the empiricists called primary and secondary qualities: the colors, shapes, and textures of ordinary things, as well as the basic properties of entities discovered by science—the mass, charge, or energy of particles. But second, we should probably also assume that predicates saying what *kind* of thing something is also express properties that are qualitative in the relevant sense: being a lion, a flatworm, a lemon, or an elementary particle, being water or aluminum, being a sofa or an internal combustion engine. (What Aristotle called “secondary substances,” as well as artifactual kinds.) With at least some of these cases, one might argue that the existence and identity of the kind are dependent on a concrete historical process and thus on particular individuals. One might argue that even if organisms with the exact DNA of tigers had evolved independently on another planet, they would not be *tigers*. If so, this kind of kind may not be purely qualitative, but at least in many cases, being an instance of a kind does not seem to be a relational property definable in terms of particular individuals, or “thisnesses.” Third, there are dispositional properties—being soluble or brittle, flexible or observable, vain or courageous. These deserve special mention since they seem to be properties that have modal consequences—consequences for what properties the individuals that instantiate them *would* have in counterfactual circumstances. The exact relationship between dispositions and properties defined in terms of counterfactuals is controversial, but it seems clear that an object’s being flexible, for example, has some conceptual connection with being such that it (that very object) would flex under certain possibly counterfactual conditions. If, as Kaplan’s anti-haecceitist holds, “there is no metaphysical reality of sameness or difference” of individuals in different possible worlds, perhaps there will be no metaphysical fact of the matter about whether a thing has a dispositional property. Fourth, there are spatio-temporal properties and relations, which as we will see play a special role in the arguments about haecceitism.

Both Adams and Lewis, in their attempt to give a general characterization of the purely qualitative properties and relations, make use of a linguistic metaphor. So Lewis:

Suppose we had a mighty language that lacked for nothing in the way of qualitative predicates, and lacked for nothing in its resources for complex infinitary constructions, but was entirely devoid of proper names for things; then the qualitative differences would be those that could be captured by descriptions in this mighty language.⁶

And Adams:

We might try to capture the idea by saying that a property is purely qualitative—a suchness—if and only if it could be expressed in a language sufficiently rich, without the aid of such referential devices as proper name, proper adjectives and verbs (such as “Leibnizian” and “pegasizes”), indexical expressions and referential uses of definite descriptions.⁷

Adams worries that this characterization of the notion might be circular (though he doubts that it is). It seems to me undeniable that it *is* circular, since its adequacy requires assuming that the primitive predicates of the language express purely qualitative properties and relations. There is no relevant *syntactic* difference between the predicates “pegasizes” and “flies,” and no semantic difference that is independent of the distinction one is trying to explain. What the linguistic characterization of qualitative properties and relations does accomplish is to put closure conditions on the purely qualitative propositions, though exactly what the closure conditions are will depend on the expressive resources of the mighty language. Adams suggests that we might want to include

⁶ Lewis 1986, 221.

⁷ Adams 1979, 7.

propositional attitude operators, along with truth functional connectives and quantifiers (so that the predicate “believes that ϕ ” will be purely qualitative if ϕ is). He does not mention modal operators. Assuming that “lawyer” and “plumber” both express purely qualitative properties, should we regard “being a lawyer who might have been a plumber” as purely qualitative?

I am suspicious of the distinction between the qualitative and the nonqualitative, which is more diverse than is sometimes supposed; I am also suspicious that the metaphor of the mighty language (like any attempt to explain a substantive distinction in terms of languages used to express it) may be misleading. Nevertheless, I will begin by assuming the distinction and by exploiting the metaphor. Later I will suggest a slightly more independent way of modeling the notion of the purely qualitative.

5. A Mighty Language

Recall our modal theory of propositions: We have a logical space, modeled by a set of points, and for each point an equivalence relation that partitions the space. One point represents the actual world, and the cells of the partition that is induced by its equivalence relation represent the maximal consistent propositions. The other points represent realizations of various counterfactual possibilities, and the partition cells that are induced by their equivalence relations represent the maximal consistent propositions that would exist if they were realized. (There is some loose talk in this informal description of what is represented, talk of merely possible realizations and merely possible propositions, but I have tried to say how the loose talk is to be traded in for a more sober account of how the model relates to what is modeled.) Now let me add a language—a standard first-order modal language and an orthodox S5 Kripkean semantics for the language, in which the points of logical space are the Kripkean possible worlds. So we need to add to our structure

a domain of possible individuals for each point. The domains may vary from point to point, though the family of equivalence relations will constrain the relationship between the domains. There are lots of details about exactly how the semantics works to be spelled out, but I am not going to worry about most of them here. I will assume that our language has singular terms as well as predicates, and in the general case I will allow that primitive singular terms may be rigid or nonrigid designators. For the purpose of exploring the task of modeling the contrasting metaphysical doctrines we are discussing, I will assume (without being completely sure about what this means) that all of the primitive predicates of our language are purely qualitative, and I will *pretend* (again without being sure what I am pretending) that our language is a mighty language, one that, to paraphrase Lewis, lacks for nothing in the way of predicates for representing any qualitative distinctions that there are to be made. We don't, at this point, want to exclude the possibility that some propositions exist contingently—that is what the bells and whistles I added to the orthodox semantics were designed to accommodate. But I will make the simplifying assumption that all purely qualitative properties and propositions exist necessarily. And while in a general semantic theory for quantified modal languages we want to allow for singular terms that are not rigid designators, this is not appropriate for the mighty language, which aims to use the distinction between predicates and singular terms to model the difference between the purely general and the particular. So I will assume that in our mighty language every individual has a rigid name and that all names are rigid designators.

Now we can define the purely qualitative propositions as those that are expressible in our mighty language with sentences that contain no singular terms. We can then consider the maximal partition induced by these propositions (which, in the general case, may be coarser than the partition induced by *all* actual propositions). Whether we are haecceitists or not, we are assuming that

the propositions determined by the qualitative partition will exist in all possible worlds.

Since our mighty language has a proper name for each individual in the domain of each possible world, we can express the thisness or haecceity of an individual with a predicate of the form $\hat{x}(x = a)$, where a is a proper name of the individual.⁸ The anti-haecceitist thesis (according to Adams's version) is that all thisnesses are reducible to, or supervenient on, suchnesses. This implies that all propositions will be supervenient on the qualitative ones. In a model that meets this anti-haecceitist supervenience condition, the equivalence relations all coincide with the identity relation. *All* propositions exist necessarily.⁹

Our first attempt to model the anti-haecceitist doctrine follows Adams's characterization, but it does not fit with Kaplan's, since our models take for granted the identification of individuals across possible worlds. And by using a mighty language that contains modal operators, we have perhaps let too much count as qualitative. Is it right to think of general claims that involve iterated modal predications as purely qualitative? Consider the property of not having any sons but possibly having seven sons, the seventh of whom is a lawyer, but a lawyer who might have not been a lawyer but a plumber instead. Or consider the proposition that some person (we aren't saying who) has this property. Is this a purely general proposition that an anti-haecceitist should admit? This may be unclear, but the anti-haecceitist will face problems if she does allow such properties and propositions into the supervenience base. Let me use the

⁸ Since the merely possible individuals are artifacts of the model (in some haecceitistic models), the names will serve only as ways of linking the possibilities of there being individuals of various kinds in different possible worlds. Sentences with such names may not actually express propositions (in the actual world). In our models, sentences express propositions if and only if their truth value stays the same under the permutations induced by the equivalence relations.

⁹ See appendix C for the technical details of a semantic model for a mighty language.

example of Kripke's dice as a very small toy model to bring out some of the problems. Just to cut our metaphysical universe down to size, suppose there are necessarily exactly two individuals—a pair of dice—and just six primitive predicates (that are disjoint and exhaustive) that say how a die lands (1 through 6). Assume that despite its small size, this is a mighty language for our toy example of modal reality. It should be true, if we have captured the intuitive idea of anti-haecceitism, that saying how each of the two dice lands will be a complete description of the universe. But this will not follow, since the model allows that one might distinguish the dice by their modal properties. Consider the sentence that says that there is an x and a y such that x landed 5 and y 6, and possibly y landed 5 and x 6.

$$(1) (\exists x)(\exists y)((Fx \wedge Sy) \wedge \Diamond(Fy \wedge Sx))$$

If (2), $(\exists x)(\exists y)(Fx \wedge Sy)$, expressed a maximal proposition, then sentence (1) could not be true in an anti-haecceitist model of the kind I described, since it would require two indiscernible but distinct possible worlds. But suppose that it is true of each of the dice that it might land either 5 or 6 but that only one of them could land 1. Then (2) is not a maximal proposition, and we would have an allegedly qualitative property (a modal property) to distinguish the actual 5-6 world from a distinct possible 6-5 world, and so sentence (1) could be true. If we allow any identification across possible worlds, our anti-haecceitist constraints will allow for modal differences between individuals that are not grounded in any of the basic qualitative properties. This seems incompatible with the spirit of anti-haecceitism.

Suppose we assume (i) that the *only* difference in the dice is the way they in fact land, (ii) that any modal properties a die might have are the same as the modal properties the other die might have, and (iii) that it is possible that the two dice land differently. Then

our anti-haecceitist constraints will require that neither of the dice could have landed differently from the way it did land. In effect, we must require that the individuals are world-bound. (So neither of the dice that landed 6 and 5 in the relevant world are literally the same as the ones that lands 1 in some other possible world.) But then if we are to come close to capturing common modal opinion, we will need a counterpart relation and a reinterpretation of our semantics for modal predicates. According to both Kaplan and Lewis, the counterpart relation is not part of our metaphysics but part of an account of how we use possible worlds to represent. Quoting Kaplan: "Our interests may cause us to *identify* individuals of distinct worlds, but then we are creating something—a trans-world continuant—of a kind different from anything given by the metaphysics."¹⁰ Modal predicates such as being a lawyer who might have been a plumber will express general properties only relative to some perhaps metaphysically arbitrary choice of a counterpart relation. Perhaps the same will be true of predicates such as being flexible or observable. If the distinction between purely qualitative and partly haecceitistic properties is (for the anti-haecceitist) a distinction between the properties that are determined by the metaphysics of the possible worlds and those that depend on the choice of a counterpart relation, then any modally loaded property will not be purely qualitative.

If the anti-haecceitist has a different semantics, as well as a different metaphysics (with modal predication interpreted with a different semantic rule), then it will be more problematic to use the mighty language strategy to characterize the difference between haecceitism and anti-haecceitism. It might be clearer, in any case, to try to model the purely qualitative in a way that is not dependent on the linguistic picture.

¹⁰ Kaplan 1975, 723.

6. Property Spaces

One way to model a structure of purely qualitative properties and relations is to use an abstract space: a quality or property space.¹¹ Qualitative intrinsic properties are represented in such a model by regions of the property space, which may have a rich structure with many dimensions. So, for example, one dimension of the space might be the color spectrum, while another is a temperature scale. The property green would be modeled by a region of the space—a set of points—determined by a segment of the color spectrum dimension. The relation “warmer than” would be modeled by a set of ordered pairs of points of the space, determined by their place on the temperature scale. In such a model, physical space might be part of what is represented in the quality space so that spatial properties and relations, in the literal sense, would also be modeled by sets of points and *n*-tuples of the quality space.¹²

Without further modification, this representation of properties and relations makes a strong metaphysical assumption: it assumes a Leibnizian principle that all relational properties are grounded in the intrinsic properties of the relata. Intrinsic properties are represented by sets of points and binary relations by sets of ordered pairs of points. So if you specify all of the intrinsic properties of two

¹¹ I discuss this kind of model in Stalnaker 1979, drawing on van Fraassen 1967.

¹² My assumption that physical space should be represented by dimensions of a property space is eccentric and controversial. Ironically, the spatial metaphor is not normally applied to space itself. It is more usual (at least on an absolute conception of space) to think of spatial points and regions as particulars that are related to objects in space. Or, in some theories, spatial points and regions are the only particulars; ordinary objects are explained in terms of the properties of spatial points and regions. Part of my puzzlement about the alleged distinction between qualitative and nonqualitative is puzzlement about how to think about spatial structure. For the Lewisian anti-haecceitist, if one thinks of absolute spatial points and regions as particulars, then one can identify spatial points across possible worlds only relative to a counterpart relation. The resulting theory looks more like a relational theory space. See Lewis 1986, 79n, for a nice discussion of alternative conceptions of space.

individuals, you determine the two points at which they are located and thus determine all the binary relations that are exemplified by the pair. This metaphysical assumption is plausible for relations of similarity and difference, and it can be maintained for spatial relations if we assume an absolute conception of space.¹³ On this conception (at least on one way of construing it), a spatial location is an intrinsic property of an object located there. Spatial location is not what comes to mind when one thinks of an intrinsic property, but (on the absolute conception of space) it is a property that an object has that is independent of the location and properties of all other objects. One might argue that even on the absolute conception of space, spatial location is relational, since it is a relation between an object and its location. This may be appropriate if one thinks of space as a substantial object, a "container" of objects. But I think it is better, even on an absolute conception of space, to think of spatial structure as a structure of properties and relations. On this way of thinking about it, to say that spatial location is relational because it is a relation to a location is like saying that being red is a relational property because to be red is to be related to a point on the color dimension of the quality space. However one models properties, a property is not relational in any reasonable sense just because to have it is to stand in the exemplification relation to it.

To reconcile the property space model of properties and relations with a rejection of this strong metaphysical assumption that all relations are grounded in intrinsic properties, we adopt the same general strategy that we use to represent any purely relational structure: add an equivalence relation and a set of permutations on

¹³ The metaphysical grounding assumption is Leibniz's, but Leibniz is also a resolute defender of a relational theory of space. How are these two doctrines to be reconciled? This combination commits Leibniz to the thesis that spatial relations, as well as spatial properties, are unreal. Leibniz seems to have recognized and accepted this consequence. The basic objects—the monads—do not stand in spatial relations at all. For Leibniz, space is a kind of ideal construction used to model the (grounded) relations between the monads.

the points of a property space that transform it into an equivalent space—one that represents the same structure of properties and relations. This is just to generalize the familiar strategy for representing a relational theory of space that we discussed in the last chapter. A real intrinsic property, on this account, will be a set of points of the property space that is invariant under all admissible permutations, and a real binary relation will be a set of pairs of points that is invariant. This addition to the model will allow that there can be real relations that are not grounded in real properties. The most prominent examples will be spatio-temporal relations. We will return to this variation on the property space representation, since it has consequences that complicate the issue about haecceitism, but for now let me stay with the simple model that satisfies the Leibnizian grounding principle.

In this kind of model, a possible world is represented by a domain of individuals and a location function that maps the individuals into the space. A complete specification of a possible world provides a domain of individuals and for each one gives a complete qualitative characterization of the individual by specifying the point in the quality space where it is located. Two possible worlds will be qualitatively indiscernible if their domains of individuals have the same cardinality, and the domains are mapped onto the same points in the quality space. (That is, worlds w and w^* are qualitatively indiscernible if and only if there is a one-to-one correspondence between D_w and D_{w^*} , and w locates a member of its domain at point x if and only if w^* locates the corresponding member of its domain at point x .) In the context of this kind of model, one anti-haecceitist doctrine is that qualitatively indiscernible worlds are identical. But this falls short of fully capturing either of the versions of anti-haecceitism we have been considering. First, on this characterization, haecceities (properties expressed by predicates of the form $\hat{x}(x = a)$) are not necessarily supervenient on qualitative properties, since we have not ruled out the possibility that

within a single world, there are two distinct individuals located at the same point in logical space. To use again our toy example of the dice, consider the possibility that both dice land 6. The two dice are qualitatively indiscernible but still distinct. So without a further constraint, we have not satisfied Adams's definition of anti-haecceitism. Second, this characterization permits irreducible modal properties of individuals, properties that are not determined by the point in logical space that an individual occupies but by the identity of the individual. Again the dice example: it might be an irreducible fact that the individual that lands 6 in the 5-6 world is the same as the one that lands 4 in the 3-4 world, or it could be instead that it was the one that lands 5 in the 5-6 world that is the one that landed 4 in the 3-4 world. So this characterization also does not meet Kaplan's requirement that there be "no metaphysical reality" to the identification of individuals in different possible worlds.

One way to strengthen the anti-haecceitist doctrine so as to satisfy both the Kaplan and the Adams characterizations is to say that we should represent a possible world not as the location of some independently specified set of individuals in the property space but simply as a specification of which points in logical space are occupied. The points in the property space are not occupied by anything—they are just occupied, or not. On this "bundle theory" of individuals, it is not just that "there is no metaphysical reality" to the identification of individuals across possible worlds—there is no metaphysical reality to the individuals themselves, even within a possible world: they are just features of a point in the property space. But it is important to see that the doctrine defended by Lewis is not this bundle theory. Lewis is, after all, a nominalist. Properties and relations are just sets of their exemplifications. It is central to Lewis's metaphysics that there is a distinction between natural properties and arbitrary groupings of individuals (without this distinction there could not be the distinction between qualitative and nonqualitative), but this is just a distinction between kinds

of sets of individuals. So Lewis could not reduce individuals to their properties without having his metaphysics swallow itself. On the bundle theory, the identity of indiscernibles (for "individuals") is a necessary truth, since an individual is just the co-instantiation of all the properties represented by a point in the space of properties, but Lewis is explicit that his brand of anti-haecceitism "does not imply any doctrine of identity of indiscernibles."¹⁴ So Lewis's doctrine is not the bundle theory and is not incompatible with the thesis that Adams identifies with haecceitism. But it does require that any modal properties be explained in terms of a counterpart relation and that the counterpart relation be specified in terms of the qualitative character of the worlds that are inhabited by the individuals that stand in the relation.

Lewis's brand of anti-haecceitism does not imply that we can read the modal properties of things off of a complete description of the possible world. I will use yet again the version of the dice example discussed earlier to illustrate this point. This time we use the counterpart version of the model. The same assumptions about the possible worlds might be combined with different counterpart relations. We could use a counterpart relation that linked the 5-die in the 5-6 world with the 4-die in the 3-4 world, or alternatively, we could use a different counterpart relation that links the 5-die in the 5-6 world to the 3-die in the 3-4 world. Both of these counterpart relations are admissible, given the version of anti-haecceitism that Lewis defends, and the choice between them makes a difference as to what modal properties the dice have. The point is that things have modal properties only relative to the choice of a counterpart relation, and the choice between alternative counterpart relations is not constrained by the metaphysics. Given a counterpart relation,

¹⁴Lewis 1986, 224. His main concern, in the context in which he says this, is with the identity of indiscernible possible worlds, but he does make the unqualified claim, and nothing in his account excludes the possibility of possible worlds with distinct but qualitatively indistinguishable individuals.

there will be an objective fact of the matter what modal properties things have, but the choice itself is conventional. If dispositional and causal properties of things are conceptually connected to the properties that those things (or their counterparts) have in other possible worlds, then things will have such properties only relative to counterpart relations that can be specified in terms of the similarities and differences between possible worlds that are characterized in terms of the noncausal qualitative character of the worlds.

Lewis, in a passage quoted above, mentions causal relations in his characterization of the qualitative. He mentioned them in order to remain neutral about the relation between causal relations and other qualitative properties: "Here I count causal relations as a matter of 'qualitative character'; it may or may not be that they supervene on qualitative character more narrowly defined, in particular on the point-by-point distribution of local qualitative character."¹⁵ But the neutrality may not be sustainable, since it is not clear that Lewis's thesis of Humean supervenience (roughly, that causal relations, counterfactual dependencies, and dispositions that hold can be reduced to the patterns of the qualitative character of things) can be separated from his anti-haecceitism. The two theses may not strictly entail each other, but it may be difficult to motivate the one without the other. Lewis's bigger picture, one that motivates both his version of anti-haecceitism and the Humean supervenience thesis, is something like this: we can and should separate the intrinsic features of possible worlds from their relational features (the similarities and differences between the worlds and the things that inhabit them). The world itself consists of a set of "world-mates" with their purely qualitative properties and relations. We do describe the world in terms of the modal properties of things and the causal and counterfactual dependencies between things, and there are objective facts of the matter about whether such descriptions

¹⁵ Lewis, 1986, 221.

are correct. But such descriptions are in principle factorable into a part that concerns the intrinsic features of the world—the purely qualitative character and arrangement of things and their world-mates—and a part that concerns the ways the world is similar to or different from other possible worlds in one or another respect. On a contrasting picture, which I think is the right one, there is no way to separate the features of a world that are characterized by relations to other worlds (cross-world identity or counterfactual/dispositional dependence) from those that are features of the world as it is in itself. I think Lewis's thesis of Humean supervenience and his doctrine of anti-haecceitism should stand or fall together.

7. The Identity of Indiscernibles

Even if Lewis is not committed to the principle of the identity of indiscernibles, Adams's version of anti-haecceitism is committed to it, and his main argument in favor of what he calls "primitive thisness" is an argument against this principle. I will conclude this chapter by looking at his arguments, which also raise some problems for Lewis's account and bring out some further connections between anti-haecceitism and issues about counterfactuals.

The basic argument, which Adams calls "the spatial dispersal argument," begins with a famous counterexample to the identity of indiscernibles presented in a dialogue by Max Black.¹⁶ Imagine a possible world that consists of just two indistinguishable iron spheres two diameters apart from each other. The composition, shape, mass, and so forth of each sphere is the same as that of the other, and they have exactly the same spatial relational properties: each is exactly two diameters from a sphere that is indistinguishable from itself.¹⁷ So, it is argued, the two spheres are indiscernible

¹⁶ Black 1952.

¹⁷ I am reminded of a line I once read in a set of directions for setting up a stereo system: "Place the two speakers equidistant from each other." Full disclosure forces

(with respect to qualitative properties and relations) but must be distinct, since each is some distance from the other.

The first question to ask is whether this is really a counterexample. The answer depends on what conception of space we are assuming. Suppose that the correct theory of physical space for the two-sphere world is an absolute theory, and that space is Euclidean. To be located at point (x,y,z) just is to have a certain property. The coordinates are conventional, but on the absolute conception, the points they name are not. One might use three arbitrary objects not all on a straight line to fix the coordinate system, but this is just reference-fixing.¹⁸ The points designated are determinate locations that we can identify across possible worlds. (We don't need counterparts here, even for a Lewisian, since we are talking about spatial properties, not particular individuals.) On this conception, *being located at a certain particular place* will be an (intrinsic) qualitative property, one that distinguishes the two spheres, so we don't have a counterexample to the principle. The spheres do share all their nonspatial intrinsic properties, as well as all of their relational spatial properties, but being in a different place is itself a spatial property, so a difference in location is enough to establish a difference that is qualitative, in the broad sense in which qualitative properties and relations are those that are not constituted by relations to particular individuals.

Adams says, in discussing a response to this alleged counterexample by Ian Hacking (a response I will consider below), "if 'absolute space-time' is accepted, the spatial and temporal dispersal arguments are quite successful in refuting the Identity of

me to admit that this was a fictional set of directions from the novel *Wittgenstein's Mistress*, by David Markson.

¹⁸I am not imagining references being fixed and a conventional coordinate system defined in the two-sphere world itself. There are no people there to do the reference-fixing (though as we will see, in one of Adams's variations on Black's world, there are). But people living in a world of the right kind could fix the references and then specify the counterfactual world of Black's spheres in terms of it.

Indiscernibles.”¹⁹ But I think this is wrong. It is only on a relational theory of space that the argument works.

Suppose we stay with the Euclidean conception of space but adopt a relational theory according to which any mapping of points onto points that preserves the distance relations between all points is an equivalent representation. On this conception, Black's example is indeed a counterexample, since all *real* qualitative intrinsic and relational properties are shared by the two spheres (where, as we said above, a real property or relation is represented by a set of points in the property space that is invariant under all admissible mappings.)

Ian Hacking responds to the counterexample in a different way, questioning not whether Black's spheres are indiscernible but whether they are distinct. “Whatever God might create,” Hacking writes, “we are clever enough to describe it in such a way that the identity of indiscernibles is preserved. This is a fact not about God but about description, space, time and the laws that we ascribe to nature.”²⁰ Adams spells out Hacking's argument, as applied to Black's example, as follows: “The most that God could create of the world imagined by Black is a globe of iron, having internal qualities Q, which can be reached by traveling two diameters in a straight line from a globe of iron having qualities Q. This possible reality can be described as two globes in Euclidean space, or as a single globe in a non-Euclidean space so tightly curved that the globe can be reached by traveling two diameters in a straight line from itself. But the difference between these two descriptions represents no difference in the way things could really be.”²¹

Hacking's argument depends on a more radical version of a relational theory of space, one that holds that less of the structure is attributed to the reality described and more is held to be an artifact

¹⁹ Adams 1979, 15.

²⁰ Hacking 1975, 255–56.

²¹ Adams 1979, 15.

of the model. To make this kind of view precise, one would have to spell out an equivalence relation or a class of permutation functions for mapping one spatial representation of a world into another, one that is held to determine an alternative representation of the same possible world. But the presupposition of Hacking's argument is a substantive empirical or metaphysical presupposition about what real differences there are between alternative ways things might be. In discussing this argument, Adams asks why the difference between the imaginary Euclidean two-globe universe and its non-Euclidean "tightly curved one-globe counterpart" can't be "a difference between possible realities in its own right? Indeed I think it is extremely plausible to regard it so."²² This seems to me to be the right response.

The upshot is that Black's scenario provides no counterexample on an absolute conception of space and, for a different reason, no counterexample on a Leibnizian relational theory according to which space as a whole is merely ideal. But it does provide a counterexample for those who accept a more moderate relational theory, according to which the relational structure of physical space is real but absolute locations are not. And I think one can give a more general abstract characterization of the conditions under which a counterexample of the form of Black's will work: the success of a counterexample of this kind will turn on whether one accepts the Leibnizian metaphysical thesis that all real relations are grounded in real intrinsic properties (a thesis that is compatible both with the absolute conception of space and with the radical rejection of the reality of spatial structure). Abstractly put, a counterexample with the form of Black's requires a case where two individuals stand in a real (symmetric) relation that requires

²² Ibid., 16. Adams adds some qualification to this judgment, since he has some sympathy for a more radical Leibnizian phenomenalism, according to which ordinary physical objects (as contrasted with minds) are not fundamental entities at all and space as a whole is merely ideal.

distinctness but nevertheless share all real intrinsic and qualitative relational properties. This will be possible only if the Leibnizian grounding principle is rejected.

In defense of the idea that we should accept the reality of the distinctness of the two indiscernible spheres, Adams considers some variations on the example.²³ Suppose that the symmetrical world is considerably more complex than Black's world, with people inhabiting two indiscernible planets, each with an indiscernible counterpart on the other planet. A person on one of the two planets might be looking through his telescope at his counterpart on the other planet, who is of course looking back at him. Each of these persons might be wondering about how the remarkable coincidence in the behavior of each twin and his counterpart is to be explained. It *might* be that there is just one person who is seeing himself, as in a mirror, but it seems implausible to assume that it is merely conventional whether this alternative description is correct. Even in the one-planet version of the story, the person, while looking through his telescope at himself, might be wondering about how the remarkable coincidence is to be explained, but it seems that in this case he would be making a factual mistake, like the mistake of the person looking at himself in a mirror and marveling at the fact that the person he is looking at is able to mimic flawlessly, instantaneously, his every move, no matter how hard he tries to fool him.²⁴ Surely we can imagine a world that is factually distinct from the mirror-like world, a world in which there really are two twins who are not making this mistake.

Notice that the judgment that there is a factual difference between the alternative spatial representations is based on judgments about different causal, explanatory, and counterfactual features of the world that are implicit in the contrasting representations. These

²³ The variations I will consider are not the same in all details as those discussed by Adams, but they are based on his examples and arguments.

²⁴ As in the classic Marx brothers routine in the movie *Duck Soup*.

considerations will have less force for the proponent of Humean supervenience, according to which such features supervene on "the point-by-point distribution of local qualitative character." The right explanatory theory is just the one that best accounts for the local facts. Lewis's specific statement of Humean supervenience presupposes a spatial structure (point-by-point distribution is a distribution in a given space), but the idea behind Humean supervenience is that the right causal-explanatory structure is whatever provides the best systematization of the particular phenomena, which is also the idea that motivates Hacking's argument about the identity of indiscernibles. But this idea does not tell us how to draw the line between the particular phenomena that provide the supervenience base and the representational structure that systematizes it.

If we accept the counterexamples to the Identity of Indiscernibles and reject the Leibnizian grounding principle, then we must reject anti-haecceitism, as Adams characterizes it—the thesis that all thisnesses are supervenient on suchnesses. But might we still accept the version of anti-haecceitism that Kaplan articulates and that Lewis defends? Could we accept primitive identity while rejecting meaningful cross-world identities that are not defined relative to a counterpart relation that is itself grounded in the qualitative similarities and differences of individuals? As I noted above, Lewis denies that his version of anti-haecceitism is committed to any version of the Identity of Indiscernibles, but Adams suggests that the two anti-haecceitist doctrines stand or fall together, and he has an ingenious argument from primitive thisness to transworld identity. I find the argument persuasive, but I think the Lewisian has some resources to resist it. Adams's argument turns on counterfactual properties of possible worlds, and the Lewisian response brings out further connections between issues about counterfactuals and issues about haecceitism.

Adams's argument begins with the assumption that in the two-spheres scenario, there are really two distinct spheres and that

they are qualitatively indiscernible throughout their history. Call the two spheres Castor and Pollux, and call the possible world of the scenario w_1 . It is assumed that the two spheres in fact maintain their existence and indiscernibility throughout their history, but, Adams argues,

it seems perfectly possible, logically and metaphysically, that either or both of them cease to exist. Let w_2 , then, be a possible world just like w_1 up to a certain time t at which in w_2 Castor ceases to exist, while Pollux goes on forever; and let w_3 be a possible world just like w_2 except that in w_3 it is Pollux that ceases to exist at t while Castor goes on forever. That the difference between w_2 and w_3 is real, and could be important, becomes vividly clear if we consider that, from the point of view of a person living on Castor before t in w_1 and having (of course) an indiscernible twin on Pollux, it can be seen as the difference between being annihilated and somebody else being annihilated instead. But there are no qualitative differences between w_2 and w_3 .²⁵

Since w_2 and w_3 are qualitatively indiscernible, the Lewisian is committed to saying that there is really just one possible world here, differently described. But the Lewisian can bite this bullet. Let me develop the response that I think he would give: "There are indeed two distinct spheres, and distinct sets of inhabitants on them, in w_1 . There are of course also two corresponding spheres in the single world that you have labeled w_2 and w_3 (I will call it w_{23}) and the spheres of w_{23} are both distinct from both of the spheres in w_1 , since individuals exist only in their own world. But both of the spheres in w_{23} are counterparts of both Castor and Pollux. All of the inhabitants of either Castor or Pollux have counterparts on *both*

²⁵ Adams 1979, 22.

of the counterpart spheres in w_{23} . (The counterparts are linked: pairs of counterparts of two *individuals* a and b are not necessarily counterparts of the pair $\langle a, b \rangle$.)²⁶ So, for example, suppose the two spheres in the variant world are called Lucky (the one that survives) and Unlucky (the one that does not); suppose further that (i) John and Mary are two inhabitants of Pollux, (ii) that Bill is a counterpart of John on Lucky, and (iii) that Susan is a counterpart of Mary on Unlucky. Then while Bill is a counterpart in w_{23} of John, and Susan is a counterpart in the same world w_{23} of Mary, it is not the case that the pair $\langle \text{Bill}, \text{Susan} \rangle$ is a counterpart of the pair $\langle \text{John}, \text{Mary} \rangle$. Now John can truly say, 'I and my sphere might have been annihilated, and it is also true that I might have survived, while Castor and all of its inhabitants were annihilated instead.' Both conjuncts of John's statement are made true by the existence of the same possible world."

The critic responds: "So there is a single world in virtue of which I might have been annihilated, and I might have not been annihilated. Doesn't this imply that I might have been both annihilated and not annihilated?" The Lewisian reply: "No, that would require that I have a counterpart in some world who is both annihilated and not annihilated, which cannot happen. Furthermore, while there is a single world in which a counterpart of John is annihilated while a counterpart of Mary is not, this does not make it true in w_1 that it might have happened that John was annihilated while Mary was not, since the counterparts in virtue of which John might have been annihilated and Mary might not have been are not appropriately linked."

To represent the interrelations between counterparts, the Lewisian's counterpart semantics should represent the counterpart relation as a set of counterpart *functions* that take a world w and a possible individual d to a member of the domain of w that is a

²⁶Lewis discusses the need for linkage between counterparts (1986, 233). The point was first made by Allen Hazen (1979).

counterpart of d in that world. The functions may be partial functions, since an individual in the domain of one world may not have any appropriate counterpart at all in the domain of another.²⁷

8. Cheap Haecceitism

There is only one possible world, w_{23} , Lewis would insist, but he acknowledges the haecceitist intuition that Adams's w_2 and w_3 are distinct possibilities. He thought he could accommodate the intuition with what he called a "cheap" version of haecceitism—one that is free of any haecceitistic metaphysics. The basic idea of Lewis's cheap haecceitism is that "*Possibilities are not always possible worlds.*"²⁸ One kind of possibility is a full possible world, but another is a possibility *for* some thing or sequence of things. Such possibilities might be represented by a pair consisting of a possible world and an admissible counterpart function; w_2 and w_3 are distinct possibilities *for* Castor and Pollux, possibilities that select different counterparts within the same world.

One further twist in the dialectic will take us to an argument against Lewis's cheap haecceitism by Delia Fara.²⁹ We have been considering the possibilities w_2 and w_3 as counterfactual alternatives to w_1 , but suppose we consider instead the possibility of w_3 from the perspective of w_2 . Don't imagine John on Pollux in w_1 contemplating the possibilities that either he and his sphere, or his twin and *his* sphere, are annihilated. Instead imagine Bill, on Lucky in w_{23} , considering the possibility (perhaps after the annihilation of

²⁷ I say an *appropriate* counterpart, since a counterpart function may take no value for a given world and individual even if the individual has a counterpart in the world. Suppose that John has a daughter Mary and that in a certain counterfactual world c , John has two counterparts, one of whom has a daughter who is a counterpart of Mary, while the other has no daughter at all. So Mary has just one counterpart in c . But the counterpart *function* that takes John to his counterpart with no daughter will be undefined for Mary.

²⁸ Lewis 1986, 230.

²⁹ Fara 2009.

Unlucky) that it was Lucky with all of its inhabitants, rather than Unlucky, that had been annihilated. "Aren't we lucky," he thinks, "that it was Lucky that was saved, rather than Unlucky." It seems that Bill, in w_{23} , is considering a possibility that is *counterfactual*, from his perspective, but the world of this counterfactual possibility is Bill's own actual world. So in what sense is it counterfactual? One of Bill's compatriots on Lucky says to him, perhaps to temper his feeling of relief, "there is really no luck involved here, for if the possibility you are envisioning had transpired, the world as it is in itself would have been no different from the way it actually is. The counterfactual possibility you are envisioning is just a different perspective on our actual reality." That does not seem quite right.

In Lewis's original exposition of counterpart theory, while he allowed that an individual might have multiple counterparts within a single possible world, he required that an individual could have only one counterpart, itself, within its own world. In later work, motivated in part by the kind of example we have been discussing, he lifted this requirement, allowing that an individual might have more than one counterpart, even within its own world.³⁰ But Fara has argued that this move not only stretches intuitive credulity but also faces a technical problem. I will conclude this chapter by describing her argument and considering how Lewis might respond to it. I think Lewis has some resources to respond to the technical problem: his cheap haecceitism can explain how a possibility that is represented by the actual world can still be a counterfactual possibility. But the problem Fara raises reemerges, even in a modified counterpart semantics that allows for a cheap version of haecceitism.

³⁰ "In 'Counterpart theory and quantified modal logic' I took it as axiomatic that nothing can have any counterpart besides itself in its own world. I would now consider that requirement appropriate under some but not all resolutions of the vagueness of the counterpart relation" (Lewis 1986, 232n).

The technical problem concerns a counterpart semantics that allows an individual to have a distinct counterpart within its world, and also has an actuality operator. I will use our example to illustrate her argument. In w_{23} , Bill (who lives on planet Lucky) is not annihilated but might have been (since he has a counterpart who is annihilated). Let ϕ be a sentence that says that Bill survives, and assume that w_{23} is the actual world of our model. Let @ be the "actually" operator. Then $(\phi \wedge \Diamond \sim \phi)$ is true in w_{23} , and since the counterpart of Bill who does not survive (at least one of them) exists in the actual world, w_{23} , it is true there that $(@ \phi \wedge \Diamond @ \sim \phi)$. But in the logic of actuality, one can infer $@ \psi$ from $\Diamond @ \psi$ (to say that it might be true that *in the actual world* that ϕ is to say that it *is* true there that ϕ). So we can conclude that $(@ \phi \wedge @ \sim \phi)$ is true in w_{23} . But if two propositions are both actually true, then so is their conjunction: we can infer $@(\phi \wedge \sim \phi)$, which is a contradiction.

The first thing I think Lewis would say in response to this argument is that he was already suspicious of the idea that "the language of boxes and diamonds affords a good regimentation of ordinary modal thought,"³¹ and this argument shows that we should be all the more suspicious of that language, supplemented with an actuality operator. On his view, the best regimentation of a theory of modal reality is as a first-order extensional theory that includes possible worlds in its domain and an indexical singular term for the actual world. In his favored formulation of counterpart theory one can state all the facts about actuality and possibility in a way that is consistent: "If the language of boxes and diamonds proves to be a clumsy instrument for talking about matters of essence and potentiality [or, he might have added, about actuality], let it go hang."³² But I want to set aside this dismissive response, which is tied to Lewis's modal realism. We are concerned not just with

³¹ Lewis 1986, 12.

³² Ibid., 12–13.

what Lewis should say, given his overall theory, but with what an actualist anti-haecceitist should say—an anti-haecceitist who aims to reconcile, in the context of our general theory of propositions, an anti-haecceitist metaphysics with the “cheap” explanation of the apparently haecceitist phenomena.

So how should the proponent of cheap haecceitism respond to Fara’s argument? I think he should say that when we distinguish possible worlds from the more general category of possibilities, we also need to distinguish the actual *world* from an actual *possibility*, and the semantics for an actuality operator should say that $@\phi$ is true relative to a given possibility if and only if ϕ is true in an actual possibility. If possibilities are represented by pairs consisting of a possible world and an admissible counterpart function, then what is an actual possibility? It is a pair consisting of the actual world and a counterpart function that takes each actual individual to itself. We then have a clear account of what it is for a possibility to be counterfactual, even if it is a possibility represented by the actual world. What this response to Fara’s argument rejects is the assumption that $@\sim\phi$ is true in a model in which w_{23} is the actual world: $\sim\phi$ is indeed true in the actual world but not in an actual possibility. In the models for this kind of semantics, the “worlds” in the intended interpretation of the relevant Kripke model are not worlds in the Lewis sense but Lewis-world/counterpart pairs.³³

But this move just postpones the problem.³⁴ Since there may be more than one actual possibility, the problem returns when we consider the actual world from the perspective of a counterfactual world. To continue with our example, suppose that w_{23} is the actual world but that we are considering what is true in w_1 . There

³³ See appendix D for a sketch of a cheap haecceitist counterpart semantics.

³⁴ Thanks to the participants in the OPW@25 conference at the University of Massachusetts, April 2011, particularly John Hawthorne, Cian Dorr, and Delia Graff, whose skeptical questions in discussion led me to look more closely at the cheap haecceitist semantics and to see that it did not succeed in solving the problem that Fara’s argument brought out.

are two relevantly different possibilities for w_1 , since there are two different counterpart functions: one taking Castor to Lucky and the other taking Castor to Unlucky. We may assume that they are both *normal* counterpart functions: they both take Castor to Castor and Pollux to Pollux within w_1 , and Lucky to Lucky and Unlucky to Unlucky within w_{23} . So there will be two distinct *actual* possibilities, one in which one of Castor's counterparts is annihilated in w_{23} , and one in which one of its counterparts is not annihilated there. It follows that it will be true (in either of those counterfactual possibilities) of Castor that he is actually annihilated and also actually not annihilated. The actuality operator, as we have defined it, is a perfectly coherent modal operator, but it does not have the right properties to be an actuality operator.

Even if one could find a semantic fix for Fara's problem, one might wonder how deep the difference is between cheap haecceitism and a more metaphysically loaded version of the doctrine. For a Lewisian modal realist, the distinction between possible worlds and more fine-grained possibilities is a big deal, metaphysically, but is there a significant difference for the actualist? We have already argued that at least in haecceitist models the points in the logical space are more fine-grained than genuine possible states of the world, which we identified with maximal propositions. But even aside from the problem about actuality operators, there remains a substantive difference, even for an actualist, between the cheap and more expensive versions of haecceitism. While a defender of primitive thisness such as Adams will agree that there are no genuine haecceitistic differences between merely possible individuals (there are no merely possible individuals, just the [generic] possibility of individuals other than those that exist), he will insist that there are facts of the matter about the identity of *actual* individuals, including facts about the counterfactual properties of actual individuals. While, on a haecceitist version of the picture I have tried to paint, there is no real distinction between two qualitatively indiscernible

universes that permute nonactual individuals (as in our dice example), but symmetrical universes that permute *actual* objects are indeed distinct, not just in the representational model but in the reality that is being modeled. As is fitting for an actualist theory, the actual state of the world has a distinctive metaphysical status, and that is a crucial difference between real and cheap haecceitism. We can *suppose* that some counterfactual world is actual, but in doing so, we are supposing not only that there are some different things, differently arranged, but also that the propositions that distinguish between genuine possibilities are different. Consider again, now from the perspective of our picture of contingently existing propositions, the three scenarios w_1 , w_2 , and w_3 with Castor and Pollux and their counterparts Lucky and Unlucky. If we pretend that w_1 is the actual world, then we can say that it is Castor itself that is Lucky in w_2 , and Pollux itself that is Unlucky in that possible state of the world; w_3 is a genuinely different state of the world, one in which Castor and Pollux are interchanged. But of course w_1 is not really the actual world, and we, in the real actual world, do not have the resources to distinguish Castor from Pollux, in w_1 . (Haven't we distinguished them by giving them different names? No, since it is not a feature of the world that they are so named; these are our labels, and there is nothing about the possible state of the world in virtue of which the names attach to one or the other.) And the point is not that there is a real distinction that we are not in a position to make (that might be true of two distant objects in our universe). It is rather that there is no feature of the actual world on which the distinction might be based, even though had that possible state of the world been realized, there would have been. Given our inability to distinguish Castor from Pollux, we also do not have the resources to distinguish the two counterfactual variants of w_1 . So while we need two points in the logical space we use to model the possibilities (w_2 and w_3), there is only one genuine possible state of the world, w_{23} .

9. Conclusion

In the end, I can offer no decisive resolution of the metaphysical issues about haecceitism. Both sides can give a formally coherent theory, and both can go some way toward accommodating the phenomena that all agree must be accommodated. What the dialectic of metaphysical argument does accomplish is, first, to connect different issues—to articulate more comprehensive contrasting pictures. The issue about haecceitism focuses on different views of the nature of qualitative properties and relations—of suchness, as contrasted with thisness. I emphasized the diversity of qualitative properties and relations, where we mean to include all properties and relations that contrast with those that depend on particular individuals, and I tried to bring out some ways that issues about haecceitism interact with questions about the nature of space and time, and about the reducibility of dispositions, counterfactuals, and causal and explanatory properties. More specifically, I argued that Lewis's thesis of Humean suervenience, his thesis of anti-haecceitism, and his modal realism are all motivated by a unified metaphysical picture that requires a sharp (and in my view problematic) distinction between the intrinsic and the relational properties of possible worlds. For the actualist, the distinctions there are to be drawn between the possible states of the world are inevitably grounded in facts about the actual world, including facts about particular individuals and about the dispositions and potentialities of things. On this picture, one can understand what a possible world would be like only by its relation to our world and to the other possible ways that a world might be.

A second thing that the dialectic of metaphysical argument might accomplish is to help disentangle semantic from substantive questions. We aim to find a more or less neutral framework in which the alternative metaphysical pictures can be sharpened and compared. There is no absolute neutrality: the aim is to find a framework that

is neutral with respect to specific controversies. The general theory of propositions I have been using is certainly not free of controversial commitments, but it aims to provide a framework that can accommodate haecceitism and its contrast, each in various forms. Some of the differences between the contrasting metaphysical accounts may in the end turn out to be notational or purely semantic, but that is a question that cannot be settled in advance. The next chapter will focus specifically on the interaction of questions about semantics and intentionality with questions of metaphysics.

Disentangling Semantics from Metaphysics

The logical empiricists held that metaphysics was cognitively meaningless. They offered a diagnosis—an explanation for the appearance of sense in metaphysics: metaphysics involved equivocation between semantic questions about what linguistic forms are most apt for describing the empirical world and empirical questions about which of the sentences of the languages we choose to speak are true, according to the rules we use to interpret those sentences. Rudolf Carnap gave a clear statement of the diagnosis in an influential paper, written near the end of the time during which logical empiricism flourished.¹ He distinguished internal from external questions; the former are questions asked within a framework that provided rules for answering the questions, while the latter are practical questions about what kind of framework we should adopt. Metaphysical questions, or pseudo-questions, arise when we confuse the two, taking the practical questions that call for a choice of framework to be factual questions about what is true. As Carnap made clear, and as is also clear in earlier statements of the positivist program, the crucial distinction is between matters of decision or stipulation and matters of discovery. For Carnap, at this late stage of logical empiricism, epistemology was holistic, and the decisions about what framework to adopt were not regarded as

¹ Carnap 1950.

arbitrary decisions but as decisions motivated by a judgment about what framework would provide a theory to best account for the phenomena. But then Quine, who shared the same holistic epistemology, argued that there is no clear behavioral difference between a decision, motivated by the facts, about what language to use and a judgment about what theoretical claim is true. The Quinean picture retains the idea that metaphysics involves mixing empirical and conceptual or semantic questions but acknowledges that there are substantive theoretical issues about how the two kinds of questions are to be distinguished. The point should not be that there is no principled difference between semantic and factual claims: rather, it should be that the distinction is not itself a distinction to be made by stipulation and not a distinction that is neutral between various alternative factual judgments. We *should* try to separate semantic from factual, but this is not something one can do at the start of an investigation before addressing substantive issues.² On the Quinean picture, metaphysics lives in the contentious border area between semantics and substantive theory.

There are different kinds of semantic questions. On the one hand, there are empirical questions about natural language semantics that interact with substantive questions about the subject matter that the language is used to talk about. On the other hand, there are questions about what kind of representational framework is best suited for modeling and clarifying the substantive issues—questions of *regimentation*. “Regimentation” was Quine’s term: for him, the practice of regimentation was similar to what the logical empiricists called “explication,” but he put a different spin on the practice. Explication was thought of as the motivated stipulation of meanings, the setting up of frameworks, with a clear semantics and well-defined rules in which the internal mathematical and

² As a textual interpretation of Quine, this may be controversial, but my concern is with the substantive point and not with whether it is what Quine meant to say. This is why I say that this is what, in my view, the point *should* be.

empirical questions could be asked and answered. Quine rejected any deep distinction between an external pragmatic decision about how to talk (motivated by the usefulness of the language for representing the phenomena) and an internal question about what is true. But he still aimed to find neutral linguistic frameworks in which controversial substantive issues about what there is, and about what can be truly said about what there is, could be framed in a clear and perspicuous way. We give up the aim of an absolutely neutral language but try to construct frameworks that are neutral with respect to specific issues. Quinean regimentation aims to avoid hard empirical questions about natural language semantics, but since it uses natural language to explain the regimented language, some issues about natural language are unavoidable. The language of quantification theory was Quine's paradigm of regimentation, and it provides both an illustration of the method and a clear case of its success. One proceeds not by trying to give a direct semantics for the quantifiers of a natural language; that is something that we still do not know how to do (or at least how it should be done remains controversial). Instead, one constructs an artificial language with a perspicuous syntax that straightforwardly reflects the compositional semantics. Scopes are marked with word order and parentheses, and cross-reference is accomplished with labeled and ordered variables. The method is not wholly free of presuppositions about the quantifiers of natural language since it is a bootstrap method that uses natural language to spell out the formal language and to give its semantics, but the hope is that one can use uncontroversially unambiguous parts of the natural language to explain the formal language and its semantics and that one can rely on competence with the natural language rather than on a theoretical understanding of that competence. And the resulting formal language can help not only to clarify one's theory of the world but also to sharpen questions of natural language semantics by giving a clear account of what distinctions there are to be made;

the empirical semanticist's job is to explain what devices are used, in natural language, to make those distinctions.

As Quine emphasized, the regimented language itself is not wholly free of commitments about the nature of the reality it is used to describe: "The quest of a simplest clearest overall pattern of canonical notation is not to be distinguished from a quest of ultimate categories, a limning of the most general traits of reality."³ To think of the world as a domain of individuals that are characterized in predications and quantified over is already to attribute a certain perhaps metaphysical structure to it. This very abstract commitment leaves a lot open and is neutral with respect to a range of very different alternative theories about what the world is like, but even in this case, there remains a mix of substantive and semantic commitments.

That is the general background. I think of the modal semantic framework we have been using as a project of regimentation of this kind. The interaction of modal semantics with modal metaphysics has been a recurrent theme in our discussions so far, but in this chapter I want to explore the dialectic of this interaction more explicitly. I will begin by looking at three examples of places where semantic and substantive issues interact, and where it is contentious whether a problem calls for a semantic or a metaphysical solution. The first example is an argument from a semantic premise to a metaphysical conclusion that I will argue is fallacious; the second example is a case of a semantic decision with a metaphysical motivation—a case where I think we can and should decouple the two kinds of issues, allowing semantic questions to be answered on semantic grounds. My third example is a puzzle where it is controversial whether the puzzle should have a semantic or a metaphysical solution. In this case I will argue that the metaphysical response is unnecessary—the puzzle is best dissolved on semantic grounds.

³ Quine 1960, 161.

After discussing these three examples, I will consider a specific semantic question about how the quantifiers should work in our modal quantification theory and the way that this question interacts with metaphysical questions about the nature of the modal reality that the language aims to describe. This will lead us, in the last section of the chapter, to a consideration of more general questions about ontological commitment.

1. Direct Reference and Object-Dependence

My first example comes from a paper by Thomas Hofwebber.⁴ His central argument is an attempt to get a metaphysical conclusion from a semantic premise. (How is that possible? Hofwebber is explicit that we need not just the premise about the semantics but also the assumption that some of the sentences to which the semantics applies are true.) The specific argument I will discuss is part of a larger argument against the thesis of physicalism: roughly, that all facts are supervenient on the physical facts, but I will be concerned only with one step in this overall argument. The semantic premise, labeled “Direct Reference,” is about the semantics for names. The metaphysical conclusion, labeled “Object-Dependence,” is essentially the thesis that Plantinga labeled “Existentialism.”

Here is Hofwebber’s characterization of his premise: “Direct reference (DR) is true if some terms in our language are directly referential. A term is directly referential if the contribution it makes to the content of an utterance in which the term occurs is merely the object referred to, as opposed to a mode of presentation, some description, or some other characterization of that object.”⁵ Object-Dependence (OD) is the thesis that some propositions (those expressed with object-dependent terms) exist only if the objects referred to exist. Here is the argument from DR to OD: “Since

⁴Hofwebber 2005.

⁵Hofwebber 2005, 7.

[according to Direct Reference] the object referred to is all that is contributed, if there is no object that is referred to, . . . then there is no complete proposition. In particular, the (complete) proposition does not exist if the object that one attempts to refer to does not exist.”⁶ Hofwebber claims that his argument is independent of any particular theory of propositions.

Now as should be clear from earlier chapters, I am prepared to sign on to the conclusion of Hofwebber’s argument, as well as to the premise, as I understand it. But I am not persuaded that the premise gives reason to accept the conclusion. Plantinga, for example, can accept DR, as I understand it, but he rejects OD, and I don’t think he can be refuted so easily. But to assess the argument, we need to say a little more about exactly how DR, stated in somewhat metaphorical terms, is to be understood. What is meant by the *contribution* made by a term to a proposition? If propositions are understood as complexes with constituents, one might take the assumption that a term contributes something to a proposition to imply that what is contributed is a constituent of the proposition. The argument may work on this interpretation, but I don’t think it is a correct interpretation of Hofwebber’s DR or that it is a plausible account of the thesis that is common to the philosophers of language who accept a direct reference account. Hofwebber claims that his thesis DR is independent of any particular theory of propositions. It should be compatible, for example, with the coarse-grained account that we have been using, an account of propositional content as truth conditions. To ensure that the thesis DR is compatible with a range of theories of propositions, I think it should be understood this way: some terms in the language refer to individuals and determine the propositions they are used to express as a function of the individual they refer to. That is, the term determines an individual, and the only job that the name does (the

⁶Ibid., 8.

only *contribution* it makes) in determining the proposition is the job of referring to that individual. The job of the rest of the expression that determines the proposition is to provide a propositional function that takes the individual determined by the directly referential expression to a proposition. Now on this understanding, the problem with the argument is this: DR implies that the determination of a proposition (via direct reference) is successful only if the individual used to determine it exists. But this does not imply that the proposition itself depends for its existence on the individual used to determine it. Suppose that no semantic property of a term other than its referent is relevant to the role of the name in determining the proposition. There is no "mode of presentation" or other mediating semantic value involved. Still, even if the proposition is determined as a function of the individual itself, the ontological status of the value of a function may be independent of the ontological status of the thing used to determine it. Just to make the point clear, consider not a *propositional* function (a function that takes a proposition as a value) but a function from persons to persons such as "mother of." The fallacy in the following argument will be evident: the contribution that the name "GWB" makes to determining the referent of the expression "the mother of GWB" is merely the object referred to by the name. So if that name fails to refer, no referent for the description will be determined. Therefore, the object that is the referent of the description is ontologically dependent on the object that is the referent of the name: Barbara Bush would not have existed if her oldest son had not existed.

Plantinga, who rejects the object-dependence of singular propositions, might agree that the determination of the haecceities that are essentially involved in singular propositions is accomplished by direct reference to objects that exemplify them. He might even agree that the *only* way to determine a singular proposition is by reference to the individual itself: perhaps we can determine a particular individual essence only as a function of an actual individual

(particular haecceities for nonexistent things might exist but be inaccessible to reference). But it is a further question whether the individual essences would exist if the individuals that instantiate them did not. As David Lewis noted, even one with *his* metaphysical views might accept the determination of propositions by direct reference to particular individuals: "So when Kripke emphatically insists that it is entirely legitimate and proper to specify worlds by making reference to individuals . . . no anti-haecceitist need hesitate to agree with him."⁷

One of the key elements of the general account of reference that Kripke and Putnam developed was that the conditions necessary for fixing reference and content are different from the conditions necessary for the existence of the referent, or the truth of the content, that is determined. So for example, while successful direct reference to Aristotle is possible only if we are causally related to him, it of course does not follow that Aristotle is ontologically dependent on being so related to us. The point is a familiar one, and I have no doubt that Hofwebber would accept its less controversial applications. But it may be debatable whether it applies in this case, since it is debatable whether propositions should be understood as structured complexes, with (in the case of singular propositions) individuals as constituents. But I think it is at least clear that Hofwebber's argument can work only with a more theoretically loaded account of propositions.

2. "Metaphysically Satisfying, though Formally Inconvenient" Semantics

The quoted phrase in this section heading comes from Robert Adams's description of the kind of modal semantics that he proposes for a language that is appropriate to the metaphysical doctrines

⁷ Lewis 1986, 222.

about modality that he favors (haecceitism about actual things but no haecceities for things that might exist but don't).⁸ I will argue that we can reconcile the satisfying metaphysics with a semantics without the inconvenient features that Adams proposes.

"There *are* no possibilities or necessities *de re* about non-actual individuals," Adams says. "The singular propositions that I exist and that I do not exist would not exist to have the logical properties . . . by virtue of which my existence or non-existence would be possible or necessary. I therefore say that ' $\Diamond(I \text{ exist})$ ', ' $\Diamond\sim(I \text{ exist})$ ', ' $\Box(I \text{ exist})$ ', and ' $\Box\sim(I \text{ exist})$ ' are all false, and their negations true, at worlds in which I do not exist. Neither my existence nor my non-existence would be possible or necessary if I did not exist."⁹

Why is a semantics that yields these judgments about modal sentences "formally inconvenient"? First, it has some radically counterintuitive consequences, which Adams points out. There are, for example, counterexamples to the basic modal principle, $\Box(\phi \rightarrow \Diamond\phi)$. On Adams's view, as on the one I have been defending, the proposition "Socrates does not exist" is true at or with respect to a possible state of the world in which the proposition that Socrates does not exist does not itself exist. But on Adams's semantics, the modal proposition "Possibly Socrates does not exist" is false at or with respect to that same world. So "Socrates does not exist is true at that world, but "Possibly Socrates does not exist" is false there.

But the more serious and systematic "inconvenience" is that Adams's theory does not give a smooth and general recursive semantics. A "convenient" semantics will give uniform rules for each operator or connective, saying how the semantic value of the complex expression is a function of the semantic values of its immediate constituents. There should be a general rule, for example, that says what the value of a sentence of the form $\Diamond\phi$ is as a function of the value of the inner sentence ϕ . But Adams's rules are like this:

⁸ Adams 1981.

⁹ Ibid., 29.

Suppose that w is a possible world and a is an individual that would not exist in w , and that $\Diamond p$ and $\Box p$ are singular propositions about a ; then $\sim \Diamond p$ and $\sim \Box p$ are true at w .

This statement is not a consequence drawn from general semantic rules for the diamond, box, and tilde but is itself a statement of a special rule for the case where the inner proposition is of a special kind. There is also a special semantic rule for quantified sentences that have a modal operator in the scope of the quantifiers, as well as names within the scope of the modal operator. But it ought to be (in a convenient semantics) that the values of multiply complex sentences (with names in the scope of operators, which are in the scope of quantifiers) are determined from the application of general rules at each stage of the recursion. It *could* be that a formally inconvenient semantics, with different rules for the same operator in different contexts, was nevertheless a well-defined semantics. But the point of regimentation is to give a perspicuous representation of the semantic structure of an expression, making it easier to evaluate the validity of arguments and to interpret complex statements. Ideally, the complexity of an idea will be shown to be the result of the interaction of component parts that are each relatively simple. Of course it is always possible that the substantive ideas that one is trying to regiment are not amenable to being broken down into simple parts, but in the case at hand, it seems to me that the formal inconvenience that Adams saddles himself with is gratuitous. The metaphysics that he wants to represent seems to me exactly right (it is the metaphysics that I have been trying to make room for in the development of the framework in the earlier chapters). The metaphysical doctrine is one that is haecceitist about actually existing things but that rejects the existence of haecceities of things that do not exist. But this metaphysics is perfectly compatible with a straightforward recursive semantics for a language that can recognize all the metaphysical distinctions that Adams wants to make. To set the stage for the clarification of the semantic

issue, I will start by looking back at an old modal distinction made by Quine.

Quine distinguished three grades of modal involvement: the first grade treats necessity and possibility as predicates of sentences. At the second grade, necessity and possibility are treated as *operators* on closed sentences. The third grade generalizes the operators so that they can operate on both open and closed sentences.¹⁰ For Quine, the most basic modal concepts applied to *sentences*, but we can make the same distinction for propositions (though the effect of the distinction is somewhat different when one takes propositions as the basic entities to which a necessity predicate is applied). At the first grade, necessity and possibility are predicates of propositions, while at the second grade we have propositional operators, which are interpreted by functions from propositions to propositions. One can make the same distinction with the extensional concepts falsity and negation: the first is a predicate of proposition, while the second is an operator that is interpreted by a function from propositions to propositions. There is a close connection between the predicates and the corresponding operators: $\Diamond\phi$ ($\sim\phi$) should be true if and only if the proposition that is the value of the sentence ϕ has the property of being possible (false). But it need not be that the sentence with the operator is a sentence that *says* (expresses the proposition that) the inner proposition is possible (or false). As we argued in chapter 2, if we make the assumption (which Adams and I both want to make) that propositions may exist contingently, then we are forced to make the distinction between inner and outer truth (which both Adams and I, along with Kit Fine, are

¹⁰ Quine 1953. Quine's argument was something like this: a crucial mistake, based on a use-mention confusion, is made in the move from the first to the second grade, but the real problems with this move do not emerge until one makes the move to the third grade. But once one has taken the first step, the second step is a natural one. The "illusion of understanding" is facilitated by the fact that the bad consequences come at a different point from the point where the mistake responsible for those consequences is made.

happy to make), and this distinction brings with it a distinction between inner and outer falsity, and inner and outer possibility. With these distinctions, complex sentences with the operators will not even be necessarily equivalent to the corresponding first grade claims. To use the notation introduced in chapter 2, it will not in general be true that $\Box(\Diamond\phi \leftrightarrow P\pi\phi)$ or that $\Box(\sim\phi \leftrightarrow F\pi\phi)$ (where the falsity predicate, F , is defined in terms of the truth predicate, as follows: $F =_{df} \hat{x} \sim Tx$).¹¹ Given these distinctions, we can distinguish two ways of regimenting the claim that if Adams did not exist, it would not then be possible that he did exist. The thesis might be (i) $\Box(\sim Ea \rightarrow \sim\Diamond Ea)$, or (ii) $\Box(\sim Ea \rightarrow \sim P\pi Ea)$. With Adams's meta-physical assumptions, but the convenient semantics, (i) is true, but (ii) is false. So we can say what he wants to say without deviating from a standard semantics with uniform compositional rules.

The S5 semantic rules for our modal language say that $v(\Diamond\phi) = W$ if $v(\phi) \neq \Lambda$, and $v(\Diamond\phi) = \Lambda$ otherwise. The proposition expressed by $\Diamond\phi$ is a *function* of the proposition expressed by ϕ , but the inner proposition (the one expressed by ϕ) is not a *constituent* of the proposition that is determined as a function of it. The proposition expressed by a complex sentence, such as $\Diamond\phi$, might exist in a possible world in which the proposition expressed by a constituent of the sentence, ϕ , did not exist. This can happen even with truth-functional operators: according to our theory, a proposition expressed by a disjunction, $\pi(\phi \vee \psi)$, might exist in a possible world in which neither of the propositions expressed by the disjuncts existed. (An example: "either Socrates was a philosopher or someone other than Socrates was a philosopher." Both of the

¹¹ For Quine, the first grade was conceptually more basic, but in the use to which I am putting the distinctions, there is no conceptual priority, so the grade level terminology is perhaps misleading. But to avoid the appearance of circularity (using the possibility and negation operators to define the predicates of possibility and falsity) we could define these predicates directly in terms of the resources of the basic theory of propositions: false propositions are those whose contradictories are true, and possible propositions are those whose unit sets are consistent.

propositions expressed by the disjuncts depend for their existence on the existence of Socrates, but the proposition expressed by the disjunction, which is the proposition that someone was a philosopher, does not.) The analogy discussed in the context of the Hofwebber argument is relevant here: the mother of GWB might exist even if GWB did not. GWB is not a *constituent* of his mother; he is only an object used, in a certain functional expression, to pick her out. But while it is obvious and uncontroversial that GWB is not a constituent of his mother, it is less obvious that the semantic values of subsentences and terms in a complex sentence are not constituents of the propositions expressed by the complex sentences. Perhaps Adams is presupposing a more fine-grained notion of proposition, which will impose more stringent constraints on the modal properties of propositions. But even presupposing a more fine-grained conception of proposition, the inconvenient semantics is still unmotivated for at least two reasons. First, our theory of propositions is a minimal one; it does not foreclose the possibility of further refinements of the notion of a proposition. Any theory of propositions that agrees that propositions *have* truth conditions will accept the intelligibility of the coarse-grained notion and of a semantics that takes these objects as the semantic values of sentences. For purposes of doing compositional semantics, it suffices to use a notion of semantic value that has the right grain to ensure that the semantic values of complex expressions are a function of the semantic values of their constituents, and our propositions are suitable for that purpose. But second, even if we presuppose a more fine-grained notion of propositions as structured complexes, the metaphysics that Adams is promoting still requires a distinction between propositions that are true *at* or with respect to a possible world and those that are true *in* a world. So he would still have the resources to express the thought that since there are no *de re* possibilities about nonexistent things, it is also true that if certain actual things did not in fact exist, there *would* be no *de re* possibilities

about them. However fine-grained propositions are taken to be, there will still be room for a distinction between claim (i) that if world w were realized, it would be true that the proposition that ϕ was possible, and claim (ii) that the proposition that ϕ is possible is true with respect to world w .

The upshot is that our more convenient semantics is perfectly compatible with Adams's metaphysics and is capable of representing all of the distinctions that he wants to make. We can and should decouple the metaphysical and semantic issues (in this case) and develop the semantics in a way that is motivated by semantic considerations.

3. A Puzzle about Existence

My third example concerns competing ways to respond to a puzzle: semantic versus metaphysical. Kit Fine¹² uses the puzzle to motivate a metaphysical distinction, but I will argue that the puzzle can be dissolved on semantic grounds. In this case, it is the metaphysics that seems to me gratuitous. The distinction Fine argues for (a distinction between "transcendental" and merely necessary truths) might be motivated on other grounds (though I am skeptical), but I don't think it can be motivated by its ability to dissolve this puzzle. The dialectic of argument in this case raises issues about modal intuition and about the contrast between regimentation and natural language semantics.

Here is the puzzle, as Fine presents it:

- (1) It is necessary that Socrates is a man.
- (2) It is possible that Socrates does not exist.
- (3) Therefore it is possible that Socrates is a man and does not exist.

¹² Fine 2005, ch. 9.

"The argument appears to be sound, with true premises and valid reasoning; and yet its conclusion is unacceptable."¹³

Fine considers two semantic solutions, both of which he rejects. I agree that these solutions, as he states them, are not adequate, but I think a semantic solution that adds nothing to the metaphysics can be defended.

The first of the semantic solutions Fine considers is attributed to Arthur Prior. The proposal is that propositions concerning an object are neither true nor false in possible worlds in which the object does not exist. According to the Priorian diagnosis, the paradoxical argument equivocates between a strong and a weak reading of necessity: strong necessity is truth in all possible worlds, while weak necessity is falsity in no possible world. Premise (1) is true only on the weak reading, while premise (2) is true only on the reading of "possible" that corresponds to the strong reading of "necessary." So the argument is invalid.

The second of the two semantic solutions Fine considers, which he calls "the standard response," assumes bivalence but distinguishes a qualified from an unqualified sense of necessity. A proposition is necessary (possible) in the unqualified sense if and only if it is true in all (some) possible worlds. A proposition "concerning certain objects" will be necessary (possible) in the qualified sense iff true in all (some) possible worlds in which those objects exist. The argument equivocates between the two senses. Premise (1) is true only in the qualified sense, while premise (2) is true in the unqualified sense.

I agree with Fine that neither of these two responses is adequate, but I still think that the puzzle should have a semantic solution. Since the solution I favor is a bivalent one, it is closer to the standard response than to the Priorian one, so I will concentrate on Fine's criticism of it. One problem for this response (one not discussed by

¹³ Ibid., 329.

Fine) is that it suffers from a problem similar to the one we raised for Adams's proposed semantic rules. If we were to spell out semantic rules for operators expressing the qualified senses of necessity and possibility, these rules would have to make reference to the internal structure of sentences on which the operators operate. Specifically, the idea is that when the internal proposition "concerns certain objects," the value of the necessitation of that proposition will depend on those objects. In the orthodox semantics, propositions expressed with rigid referring expressions will be determined as a function of the individuals to which those expressions refer, but the individuals will not be constituents of the proposition determined. Giving compositional rules for Fine's qualified senses of necessity and possibility would require finer-grained propositions and new rules for all of the connectives and operators to ensure that the new fine-grained values for the sentences were a function of the values of their components.

Even aside from this general structural problem, there are problems with the standard solution. First, Fine argues that the explanation of the qualified sense of necessity seems to imply that it is necessary (in this sense) that Socrates exists. But it seems implausible to suggest that there is any sense of necessity in which objects such as Socrates (whose existence is contingent in the unqualified sense) exist necessarily. I think this is a persuasive objection.

Fine's second point against the standard response is that it "fails to provide an adequate reason for supposing that the first premiss [that it is necessary that Socrates is a man] is not to be accepted, under the unqualified reading of 'necessity.'" This objection (Fine argues) does not apply to the Priorian response since, as he puts it, the Priorian "has a general horror of non-existence" while "the standard response has only a selective horror of non-existence."¹⁴ I don't know that for either the Priorian or the standard response

¹⁴ Ibid., 331.

we need the rhetoric suggesting a kind of irrational metaphysical phobia, but the point is that for the Priorian, *all* sentences containing nondenoting names, without qualification, will fail to be true (by being neither true nor false), while for the standard response, which is bivalent, *some* sentences with names that are nondenoting in a certain possible world will be true at that world, and so even though Socrates exists only contingently, there are (according to the standard response) truths about him that are necessary in the unqualified sense: "It will be an unqualified necessity, for example, that Socrates exists or does not exist or that possibly Socrates is a man. But this means that if we are to reject the unqualified necessity that Socrates is a man, then it must be because of some special feature of the term 'man.'"¹⁵ Fine's argument here is that since the standard response allows for *some* predicates that are not existence-entailing, it needs an explanation for why the predicate "man" is one of the ones that *is* existence-entailing. Perhaps there is some notion of a *basic* property and a thesis that all basic properties are existence-entailing. But (he argues) a property such as being a man presumably has a (real) definition, and what ground is there to assume that the definition in terms of "basic" properties will show the defined property to be existence-entailing?

I think this objection is misconceived, since in a proper formulation of a bivalent semantics that is appropriate for the standard response, the horror of nonexistence (if that is what it is) is as general as it is with the Priorian response and is not dependent on any metaphysical distinction between basic and non-basic properties or predicates. The relevant distinction is an abstract structural distinction between predicates and sentences, not a substantive or metaphysical distinction between kinds of predicates or properties. In a natural and perfectly general semantics for a bivalent language, *all* predication sentences are existence-entailing and thus false

¹⁵ Ibid..

when any of the singular terms to which the predicate is applied fail to have a denotation. We may, in such a language, have complex as well as simple predicates, but if our semantics allows for singular terms that fail to denote in some possible world, then we must distinguish predications with complex predicates from complex sentences with parts that are predications. For example, we need to distinguish the claim that it is not the case that the king of France is mortal (which is true) from the claim that the king of France is immortal (which is false). The distinction is not based on a distinction between the predicates "mortal" and "immortal" but on the fact that the "it is not the case that" is a sentence operator: the claim that it is not the case that the king of France is mortal is not the application of a predicate ("It is not the case that ____ is mortal") to the alleged king but is the denial of a sentence that purports to ascribe a property. So the truth and falsity judgments are the same if we interchange "mortal" and "immortal" in the pair of sentences. "Mortal" and "immortal" are interdefinable, but we need no assumption that one is more basic than the other to distinguish "t is mortal" from "not: t is immortal," and "t is immortal" from "not: t is mortal." Fine's examples of statements about Socrates that are unqualified necessities are cases of complex sentences, not predications. ("Socrates exists or does not exist" and "possibly Socrates is a man.") These are indeed necessary truths, but if we were to consider instead corresponding predications, with complex predicates, the claims would not be necessary (in the unqualified sense) since they would be existence-entailing. If one first forms the complex predicates "either exists or does not exist" and "is possibly a man" and then applies them to Socrates, the resulting sentences will be false in possible worlds in which Socrates fails to exist.

In the usual formulation of first-order quantification theory, either extensional or modal, there are no explicit complex predicates, though complex predicates are implicit in quantificational contexts. To say that everyone is either male or female is implicitly to say

that the predicate, being either male or female, applies to everyone. In extensional semantics, with no quantifiers other than "all" and "some" and with the presupposition that the domain is nonempty and that all singular terms refer, one needs complex predicates only for one purpose: to express universal and existential generalization. So one can get by with a language that combines two conceptually distinct semantic operations (complex predicate formation and claims of existence and universality) in one fused operator, a variable-binding quantifier. But when the language is enriched in various ways, and restrictive existential presuppositions are relaxed, there are other uses for complex predicates, and it is in any case conceptually clearer and more perspicuous to separate the two very different jobs that are fused in the quantifiers. We should have one (variable-binding) operator that forms a complex predicate from an open sentence that expresses a propositional function, and a separate device for expressing generality—an operator for saying that the predicate it operates on applies to everything, or to something, in the domain. If we have a device for forming complex predicates (which we need anyway), then variable-binding need be involved only in complex predicate formation. If we have a predicate "is either male or female" ($\hat{x}(Mx \vee Fx)$), then we can say that everyone satisfies this predicate simply by having an operator "everyone" to which that predicate is applied.

A Tarskian semantics of the traditional kind, with variable-binding quantifiers and no complex predicates, can use quantifiers together with identity to make the distinctions that complex predicates make explicitly. So we can simulate the claim that t is non-mortal, $\hat{x}\sim Mx(t)$, with the existential sentence $\exists x(\sim Mx \wedge x = t)$. One uses a quantifier designed to express generality and then uses identity to subtract the generality from the claim. This kludge works, but it distorts the conceptual structure of the complex predication. But whether one agrees with me or not that it is conceptually clearer to formulate predicate logic with explicit complex

predicates, one should agree that the issue is a purely semantic one, since the semantics for a language with complex predicates uses exactly the same resources as the classical theory. One might question the semantic decision, made in the context of this formulation, to treat predications as uniformly false when one or more of the singular terms are nondenoting, but it is a coherent decision that requires no problematic distinctions or substantive metaphysical assumptions.

In the context of this kind of semantics, the right diagnosis of the puzzle is not (as in Fine's standard response) based on a distinction between a qualified and an unqualified necessity operator, both operating on sentences. We need only the unqualified necessity operator, but we can define a complex predicate in terms of it, for any predicate F , that expresses the property of being essentially F : something is essentially F if and only if it is such that necessarily, if it exists, it is F : $(\hat{x}\Box(Ex \rightarrow Fx))$. The right semantic response to the puzzle, I think, is simply to reject premise (1) of the problematic argument. It gets its appearance of plausibility from the fact that there is something in the vicinity that is true—namely that Socrates is essentially a man, which we regiment as follows: $\hat{x}\Box(Ex \rightarrow Mx)(s)$.

I think this is the right response to the puzzle but more needs to be said in its defense. For one thing, this response suffers from a problem analogous to one that Fine raises in criticizing the standard response. On this account of "essential," we must say that Socrates essentially exists, since he has the property of being such that necessarily, if he exists, then he exists. But this seems obviously wrong. What is to be said about this objection? Different moves made in the attempt to counter it will help clarify the project of regimentation and its relation to questions of pragmatics and empirical semantics. I will first look at a dismissive response to the objection and then briefly at a response that takes it more seriously.

The dismissive line is this: The task of regimentation is to construct a language, and a semantics for it, that is adequate for saying in a perspicuous and unambiguous way what we think there is to be said. We have to use our natural language to do the semantics for the regimented language, but our hope is that we can rely on relatively uncontroversial natural language constructions, staying away from contentious terms and problematic scope phenomena. Once we have our regimented language, then as long as its semantics is clear, we can address the substantive questions in it. Now it may be unclear or controversial exactly how the semantics for words like "necessary," "possible," "essentially," "might," and "must" should go in natural language. It is clear that these words interact with each other and with quantifiers and conditionals in complex and not fully understood ways. Our diagnosis of the puzzle does not commit itself on those questions. In particular, the regimentation of "Socrates is essentially human" is not a general proposal about how the English word "essentially" works. But we have a pretty clear grasp of a regimented language for talking about modal phenomena, and it allows us to separate clearly questions about what the modal facts are from questions about how they are described. The semantics for our regimented modal language (particularly the decision that all predication sentences should be existence-entailing) does force us to say some counterintuitive things. That is, the semantics makes true some statements whose *prima facie* most natural paraphrases in natural language seem, intuitively, to be false. For example, the semantics implies that if there are contingently existing things, then it is not true that everything is necessarily self-identical. (Though it will be true that everything is essentially self-identical and that necessarily everything is self-identical.) But once it is clear (from the semantics of the regimented language) what such claims mean, it is easy to see that they are correct and unproblematic.

I have some sympathy with this dismissive response; I think it is right to locate the problem in the semantics and pragmatics rather than in the metaphysics. But the task is to accommodate certain modal intuitions, expressed in our ordinary pre-theoretic modal language, and it can't just be assumed that the source of the tensions that the intuitions reveal is in the language rather than in the substantive phenomena. If further substantive distinctions available in a richer regimented framework could be defended, and if they made it possible to resolve the tensions, that would be a reason in favor of a metaphysical response. The alternative is to say how the facts about the use of natural language might explain the conflicting intuitions. I will look briefly first at some of the resources available for a semantic/pragmatic explanation and second at the metaphysical distinction that Fine thinks can help resolve the puzzle.

Consider, as an analogy, the following less puzzling puzzle: It is necessary that all humans are animals. From all humans are animals, it follows logically that some humans are animals, which is logically equivalent to saying that some animals are humans. What follows by logic from a necessity must itself be necessary, so it follows that it is necessary that some animals are humans. But this is obviously not necessary, since evolution might have taken a different course, one in which there were animals but no humans.

This is not a puzzle that will trouble a contemporary philosopher who was raised on the regimented language of quantification theory. A dismissive response seems appropriate. Once one clarifies what is meant by a universal generalization such as "All humans are animals," it becomes obvious that the corresponding existential generalization does not follow. (Although some may still be tempted to infer from the unqualified "everything satisfies ϕ " to "something satisfies ϕ .") Still, the all-to-some inference was accepted in many traditional logics, and one needs an explanation for its intuitive pull. An explanation, or at least the rough idea of

one, is not hard to find: in normal contexts, one *presupposes* that the subject term in quantified generalizations is nonempty, and it is natural to take inferences to be logically compelling when the conclusion follows from the premise, together with what is being presupposed.¹⁶ The constraints on (pragmatic) existential presuppositions will be different when there is explicit mention of existence. It is a familiar observation that to say that tame tigers exist is not naturally understood as predicating existence of the tame tigers. It would be distinctly odd to say something like "all humans exist," and it is not just that this is trivially true, since it is odd in a way that "all humans are humans" is not. The assertion "all humans exist" is odd because it makes conflicting demands on what must be presupposed. So there is plenty of precedent for regimenting claims differently when the predicate involved is the existence predicate. It seems natural to paraphrase the claim that Socrates is essentially human as the claim that nothing could be Socrates if it was not human, but not natural to take the claim that Socrates essentially exists (not a very natural claim in any case) as the (trivial) claim that nothing could be Socrates if it did not exist. Much more needs to be said about pragmatic presuppositions, their role in reasoning, and the way they interact with modals, predication, and existence claims if one is to give a detailed semantic/pragmatic explanation for the intuitive phenomena that lie behind the tensions in Fine's puzzle. These are interesting questions in their own right, but it seems to me plausible to think that answering them can explain the phenomena without complicating our metaphysics. But we should consider Fine's alternative solution.

The response to the problem that Fine favors appeals to a distinction between two kinds of necessary truths: *transcendent*

¹⁶The notion of presupposition I am using here is purely pragmatic. One presupposes that ϕ in a given context of one is taking it to be common ground in the context that ϕ . Nothing is implied about the truth value of what is said in case what is presupposed is false.

necessities and *worldly* necessities. The distinction is explained by an analogy with the traditional distinction between sempiternal and eternal truths (propositions that are true at all times, and those that are true independently of time). The idea of the analogous modal distinction seems to be that some necessities are true in all possible worlds, while others are true independently of the possible worlds. I confess that I have only a limited grasp of the temporal distinction and no intuitive sense at all for the analogous modal distinction. I am inclined to think of the temporal distinction in modal terms: with some facts that obtain relative to all times, one can imagine that they did not, as with the spheres in the examples discussed in the last chapter, which persisted forever but were annihilated in counterfactual possible situations. With eternal truths, the truth relative to all times is necessary. But this of course won't help us grasp the analogous modal distinction. It is not that certain necessities just happen to be necessary, while others must be.¹⁷ Another way to try to get a grip on the distinction is to think about what *grounds* various truths: in virtue of what is a certain necessary truth necessary? Some are grounded in facts about the worlds, while others are true in virtue of something independently of the worlds. The fact that Socrates is either snub-nosed or he is not snub-nosed is (according to Fine) a clear case of a worldly necessity, since it is grounded in one or another contingent fact: in some worlds, the fact that Socrates is snub-nosed, and in others the fact that it is not the case that Socrates is snub-nosed. But I don't know how to apply this distinction more generally. Fine takes the propositions that Socrates is human and that Socrates is self-identical to be transcendent necessities, but why are they not necessities that are grounded in, or true in virtue of, Socrates' existence? If Socrates did not exist, it seems reasonable to say, there would be no fact that *he* was human or self-identical. So I don't understand

¹⁷ One can make sense of necessary versus contingent necessities in a non-S5 modal semantics, but this will not be the distinction Fine has in mind.

the distinction, but I also don't understand how, if we accept it, it solves the problem. Fine's diagnosis is complicated, requiring decisions about how to interpret statements that combine worldly and transcendent modality, but why can't we bypass all that and run the argument with a necessity that is agreed to be a worldly necessity?

- (1) It is necessary that Socrates either is snub-nosed or is not snub-nosed.
- (2) It is possible that Socrates does not exist.
- (3) Therefore it is possible that Socrates does not exist and either is snub-nosed or is not snub-nosed.

This seems to me exactly as puzzling and problematic as the original argument, but nothing but worldly necessities and possibilities are involved in the premises and conclusion. I don't see how the distinction between two kinds of necessity can't help with this version of the puzzle, but it seems to be essentially the same puzzle.

4. Proxy Actualism and the Range of the Quantifier

The final thing I want to do in this chapter is to consider the relation between metaphysical questions about possibilia and semantic questions about the way quantifiers are used and interpreted. This will take us to more general questions about ontological commitment.

The general line that I have been defending in this book is that an actualist who holds that there might have existed things that do not in fact exist can still do the compositional semantics for a (regimented) modal language in the orthodox Kripkean way. I argued further that one could do this in a way that does not make metaphysical commitments either to primitive haecceities for non-existent things or to an anti-haecceitist reduction of individual essences to necessarily existing qualitative properties but that this required a certain interpretation of the Kripke models, one that treated the

points in the logical spaces of the models in a way that was analogous to the way that a relational theory of space treated the locations in physical space. As Karen Bennett has said, “no actualist can take their Kripke semantics straight up; any actualist who wants to make some room for the genuine possibility of aliens [the possibility of the existence of things that do not in fact exist] will need to issue some kind of modification or caveat.”¹⁸ But whether one’s way of reconciling orthodox Kripkean semantics with actualism involves extra metaphysical commitments or not, the modification or caveat will be in the interpretation of the semantics, and the hope is that we can separate the technical questions about how to do the semantics from the metaphysical questions involved in interpreting it. But the separation is not so simple, since interpreting the semantics itself involves a mix of semantic and metaphysical commitments. We must say what we mean by the modal expressions in terms of the resources we think the world offers us for interpretation.

We can see some of the interaction of semantic with metaphysical decisions by starting with a question about how to do the compositional semantics that looks to be internal to the Kripkean semantics and independent of its interpretation. However they are interpreted, the Kripke model structures themselves make use of a domain of all possible individuals, a domain that is the union of the domains of all the possible worlds, and since the domain of some nonactual possible worlds may contain things not in the domain of the actual world, this domain of all possible individuals will, or may, be a proper superset of the domain of what actually exists. Whatever one’s strategy for reconciling this feature of the semantics with actualism, we can ask the following purely semantic question: Should the quantifier of our language range over the big set, $U = \bigcup \{D_w : w \in W\}$, or should it range over the subsets of

¹⁸ Bennett 2005.

U in terms of which it is defined: the domains of the individual worlds, D_w ? Since we are taking the basic variable-binding operator to be a predicate-forming operator, the question is really about how our semantics should define complex predicates. The general form of the semantic rule for complex predicates (using the standard Tarskian device) is this:¹⁹

$$\nu_w^s(\hat{x}\phi) = \{d \in __ : V^{s(d/x)}_w(\phi) = 1\}.$$

The question is what to put in the blank. To quantify over all *possible* individuals, we need to put in U, while to quantify over just the *actual* things (the things in the domain of the world in question), we need to put in D_w . So the two alternative rules are

$$(U) \nu_w^s(\hat{x}\phi) = \{d \in U : V^{s(d/x)}_w(\phi) = 1\}$$

$$(A) \nu_w^s(\hat{x}\phi) = \{d \in D_w : V^{s(d/x)}_w(\phi) = 1\}$$

The corresponding rules for quantifiers will be as follows:

$$(QU) \nu_w^s(\forall F) = 1 \text{ iff } \nu_w^s(F) = U.$$

$$(QA) \nu_w^s(\forall F) = 1 \text{ iff } \nu_w^s(F) = D_w.$$

Note that one cannot combine (A) with (QU), since then all universal quantifications would be false in any world whose domain is a proper subset of U. One could combine (U) with (QA), though it would be an unmotivated restriction on expressive power. One might, if one chose U, have two different universal quantifiers, one

¹⁹ "s" is a denumerable sequence of possible individuals, determining an assignment of values to the variables: if x is the n th variable in the alphabetic order of variables, then $s(x)$ is the n th term in the sequence. $s(x/d)$ is a variant sequence, defined as follows: $s(x/d)(x) = d$, and for all $y \neq x$, $s(x/d)(y) = s(y)$.

interpreted by (QU) and the other by (QA), though if one had a predicate of existence, the A-quantifier could be defined in terms of the U-quantifier plus the predicate of existence. Thus it seems reasonable to take the choice to be between the package of (U) and (QU) or the package of (A) and (QA).

Whatever one's story about interpretation, either decision is possible, and the different decisions do not seem to involve any difference in one's commitments. Whichever choice one makes, the Tarski-style semantics will appeal to a sequence of members of U and so will involve quantification over U in the metalanguage. Do the differences between different strategies for reconciling the semantics with actualism motivate one or the other of these semantic decisions? I will look at the decision from the perspective of two strategies that I have contrasted with my own, both of which are versions of what Karen Bennett has labeled "proxy actualism." The first is Plantinga's interpretation of the domains as domains of primitive individual essences, and the second is the Williamsonian metaphysics, also defended in a paper by Bernard Linsky and Edward Zalta.²⁰ On this view, as we have seen, everything that might exist does exist, and everything exists necessarily, but there is still a real distinction underlying the apparent contingency of the existence of persons and physical objects. It is contingent that such things should exist in space and time—should have the contingent property of being *concrete* objects. What we are inclined to label "merely possible things" are things that lack concreteness.

Bennett calls both of these theories "proxy actualism" because each populates the domain of the actual world²¹ with proxies for what we are inclined in our loose talk to call merely possible things—things that don't exist but would if the world were different

²⁰ Williamson 2001; Linsky and Zalta 1994.

²¹ Bennett focuses on Linsky and Zalta rather than Williamson, since Linsky and Zalta claim that their theory is a kind of actualism, while Williamson avoids saying this. But whether one calls it actualism or not, the view is essentially the same.

in various ways. Bennett minimizes the difference between the two versions of proxy actualism, but I think the difference is important. In particular, I will argue that they motivate contrasting answers to our semantic question.

First, Plantinga's theory. The domains in Plantinga's interpretation of Kripkean semantics are sets of essences, and the values of variables are essences. For the abstract semantics to work, the values of predicates must be functions from possible worlds to sets of essences (for one-place predicates; more generally, the values of relational predicates will be functions from worlds to sets of n -tuples of essences). But it would be a mistake to think that predicates in a semantics for the Plantingan theory express properties of essences. Intuitively, a predicate F will stand for a property such as being snub-nosed; the "extension" of the predicate, in a technical sense, will not be the set of people who are snub-nosed but the set of essences such that the instantiation of the essence is snub-nosed. The semantic value (relative to a given possible world) must be the set of essences and not the set of instantiations, since relative to some possible worlds, the "extension" will include essences that are actually uninstantiated, and so for the actualist, there will be no instantiations to be the semantic values, relative to those worlds. This is not a problem for the Plantinga theory—on the contrary, it is something that Plantinga would emphasize—but it does imply that the relation between predicates and the properties they are intended to express is slightly different than it is in the simple theory. One will have to be careful to distinguish two senses of "extension": the extension of a predicate, in the ordinary sense, is the set of individuals that instantiate the property expressed, while the extension in the technical sense (the semantic value of the predicate, relative to a possible world) will be the corresponding set of essences. The distinction has two consequences. First, while in the abstract formal semantics one can make perfectly good sense of a function from possible worlds to sets of essences that include essences that

are uninstantiated in that world, these functions will not represent properties. (That is, suppose for a one-place predicate F , $v_w(F) = \Gamma$ and for some essence $d \in \Gamma$, $d \notin D_w$. Then the value of F will not properly represent any property, since there will be something in its technical "extension," relative to some world, that would not correspond to anything that would [if that world were realized] be in its [ordinary] extension.) To keep the intended interpretation in line with the technical semantics, one needs to require (as a Plantingan does) that the values of predicates, relative to any world, be members of the domain of that world. This requirement should apply to complex as well as simple predicates, which means that our variable-binding rule must be (A) rather than (U), and so our rule for the quantifier must be (QA) rather than (QU). And this is the decision Plantinga makes. The second consequence of the distinction between the technical and the ordinary sense of extension is this: For the Plantingan, essences are entities in their own right and so will have properties that are different from the properties of the things that instantiate them. (So for example, an individual essence of Socrates has the property of being an individual essence of a contingent thing, and the property of existing necessarily.) To account for such entities, Plantinga will need individual essences of individual essences, which will be distinct from the essences themselves. The individual essence of Socrates has (in the actual world) the property of being a necessarily existing essence of a contingent thing, while neither Socrates himself nor the individual essence of Socrates' individual essence has this property. I see no way of avoiding an infinite hierarchy of higher-order individual essences. I don't see a technical problem with such a hierarchy, but it does perhaps strengthen the feeling of excessive metaphysical extravagance.²²

²² I cannot put much weight on the implausibility of this hierarchy since I myself am happy to accept higher-order properties and relations. There is the property of being Socrates, definable in terms of identity and the individual, and the property of being the property of being Socrates, definable in the same way in terms of identity and that property, and so on. The hierarchy that Plantinga requires seems intuitively

The situation is different on both counts for the Linsky/Zalta (LZ) version of proxy actualism. Strictly speaking, there is no difference between (A) and (U) in that theory, since the metaphysical assumption is that all worlds have the same domain: $U = D_w$ for each w . But LZ want to connect their account with ordinary modal opinion, so they have a predicate in their language that picks out a subset of the domain that are the things that "exist" according to the vulgar, and the extension of this predicate will vary from world to world. They use the word "concrete" for this, but that seems too restrictive, since it excludes abstract objects such as numbers and properties. But however it is expressed, suppose we have predicate that excludes things that could have existed in space and time but do not. Call the set of things that are in the extension of this predicate in world w ' D_w '. There is no reason why we could not have a restricted quantifier that ranges over just the things in D_w , but it would be gratuitously restrictive if it were our only quantifier (or if our variable-binding rule defined complex predicates relative to the restriction to D_w). If we have the more inclusive U rules (for both complex predicates and quantifiers) that range over everything there is, we can still capture the restricted quantifier in the standard way. If the vulgar-existence predicate in question is V , and the restricted quantifier is $\forall_v$, then the restricted universal quantifier can be defined like this: $\forall_v F =_{df} \forall x (Vx \rightarrow Fx)$.

So despite some formal parallels between the two versions of proxy actualism, it seems that one strongly motivates quantifiers that exclude what corresponds to mere possibilities, while the other strongly motivates quantifiers that include them.

It should also be noted that for LZ, there is no need for anything corresponding to the hierarchy of essences required by Plantinga's theory. The entities that are proxies for what (according to

more problematic since his haecceities are prior to and independent of the individuals that instantiate them, but I have no account to offer of what this priority and independence come to, beyond the fact that, in the case of contingent objects, the haecceities can exist even when the individuals do not.

the vulgar) are mere possibilities are the very individuals who are people or physical objects in other possible worlds. To accept the thesis that Socrates himself (and not just his essence) exists in possible worlds in which he was never born or lived is to bite one kind of metaphysical bullet, one that the Plantingan need not bite. But to accept instead an infinite hierarchy of individual essences of individual essences, all for possible individuals that do not exist (to speak loosely), is to bite a metaphysical bullet of another kind, one that the LZ theorist can avoid. I would rather dodge both bullets, but I think these metaphysical alternatives are more different from each other than Bennett suggests.

But now let me consider a different interpretation of the LZ theory, one that separates, in general, ontological commitment from the range of the quantifier. Graham Priest is one philosopher who argues for an ontologically noncommittal use of the quantifier; Nathan Salmon also defends a version of this view. Call such a philosopher a neutralist. The neutralist suggests that we drop the qualifier "vulgar" in our characterization of the vulgar-existence predicate *V* and just call it by the name "existence." He then adds that our unrestricted quantifier is existentially noncommittal. The Quinean complains, "You are ruining the good old word 'exist' . . . okay, you can have it, but now I have to find another way to describe your ontological commitments." The neutralist responds that the evidence about the use of the word "exist" in natural language is not clearly on the Quinean's side, but we should not tie the issue to a question about the meaning of the word "exist" in natural language. (Quine would surely not want to rest his thesis on the claim that it is *analytic* that to exist is to be the value of a bound variable.) A proper interpretation of the issue, I think, takes us back to the Carnapian roots of Quine's views about ontology. For Carnap (and here I think Quine would agree), if you accept a framework that involves the full apparatus of standard first-order quantification and thereby commit yourself to the intelligibility of

questions about the extent of the domain, about which predicates of one's language are correctly predicated of which members of the domain, and about questions of identity and distinctness of what is picked out by various specifications of members of the domain, then that is all there is to ontological commitment to the domain. Further *ontological* distinctions are illegitimate metaphysics. One may allow for various kinds of distinctions between members of the domain, including very general and abstract distinctions between categories of objects, but they require a kind of explanation not required for an explanation of the concept of existence. You may use the word "existence" as you like, Quine might say (and we set aside the question of whether linguistic precedent justifies you in using it as you do), but if "exist," as you use the term, picks out some subset of the domain of the quantifier then you need to specify what it means—what the criteria are for existence. The philosophical mistake is to reject the need for such specification by relying, implicitly, on an abstract explanation of the concept of the kind that Quine relies on (namely on the assumption that what exists is what is in the domain that defines the range of the quantifier, and that judgments about what is in that domain should be made on pragmatic/theoretical grounds), while at the same time taking the quantifiers to range over a domain that is larger than the class of what exists. The dialectical emphasis in Quine's methodology of existential commitment should be not on his realism about ontology but on the Carnapian rejection of metaphysical distinctions that can be divorced from the requirement of pragmatic and empirical justification.

Nathan Salmon is a neutralist who has used temporal examples to make the point that we may refer to, and quantify over, things that don't exist.²³ Socrates, for example, does not (any more) exist, but we can still talk about him. But is Salmon's point essentially a

²³ Salmon 1987, 1998.

point about the proper use of the word "exist," or is it a substantive metaphysical point? Is, for example, his denial of existential commitment to Socrates tied to a kind of presentism? It is true that Socrates was a philosopher. What are the facts that make this true? The presentist, as I understand that doctrine, takes on the obligation to explain, in terms of what exists (at the present time), what it is that makes statements about the past true. But consider a theorist who agrees with the presentist that Socrates no longer exists, and so is not a present thing whose properties (past or present) can ground the truth of the claim that Socrates was a philosopher, but argues that we don't need to ground truths about the past in presently existing entities. The present truth of the claim that Socrates was a philosopher, this theorist holds, is grounded in a fact about the past, the fact that Socrates (that very person) once engaged in the practice of philosophy. That is not presentism, and I don't think Salmon presents his thesis as a version of presentism.

Now Salmon wants his domain to include *some* things that never existed and never will exist, such as his famous character Noman, a person who would have existed if a certain (actual) pair of sperm and egg had, contrary to fact, gotten together to produce a child. If he combines this thesis with actualism (an analogue of presentism), then I think he will take on the obligation to explain truths about such nonexistent things in terms of the actual properties of actual things. (Where actual things include nonexistent things that did or will actually exist.) It is not hard to see, in cases like Noman, how such an explanation might go. The presentist will have a much harder job explaining how to ground truths about Socrates in the present.

So assuming that Salmon is an actualist, but not a presentist, I think there will be a difference in the ontological status of two different kinds of things that are, on his view, nonexistent. For Salmon (if I am interpreting his metaphysical views correctly), there should be a domain that is larger than the domain of what exists but

smaller than the whole domain of the quantifier, a domain that has a certain distinctive ontological status: it is the class of entities such that the truth of all true propositions is grounded in the properties of those objects. The Quinean claim is that our ontological commitments are to the things in the domain with this status.

LZ, I assume, take on no obligation to ground all truths in truths about the objects that satisfy the predicate *V*, and they openly acknowledge an ontological commitment to things that do not satisfy it. Like Lewis's modal realism, their view is compatible with the Quinean methodology of ontological commitment. With Plantinga, the situation is more complicated, and here is where I need to qualify my earlier claim that the quantifier in a modal language that is appropriate to Plantinga's metaphysics must be restricted to the domain of (essences of) existent things. For as we acknowledged, there would be no *formal* problem if Plantinga were to opt for the choice of *U* in his semantics. Suppose one said that all *basic* properties and relations were exemplified (in a given possible world) only by what exists relative to that world, but that complex predicates and relations could be defined with the rule *U* rather than with the rule *A*. What is expressed by such predicates are properties and relations only in an extended sense, and would apply to merely possible things. One might regard this move as a purely semantic one, claiming that the resulting theory is *metaphysically* indistinguishable from Plantinga's, even if it made a different choice from Plantinga's about how to do the modal semantics. The theory would still be an *actualist* one, if actualism is understood as the thesis that all facts must be grounded in facts about actually existing things. But one might worry, as many of Plantinga's critics have, that the metaphysics is actualist in name only, since it is achieved only by populating the actual world with entities whose nature is explained in terms of the merely possible things that would exemplify them if anything did. To continue the analogy between actualism and presentism, consider the following thesis: all truths about the past

and future are grounded in truths about presently existing things, but we can explain the past truth that Socrates was a philosopher this way: while Socrates no longer exists, his essence still does, and it is a present property of *it* that it once was exemplified by a person who engaged in philosophy. Would this count as a legitimate form of presentism? I am not sure, but it does seem a bit like cheating.

Finally, let me consider the semantic choice about the range of the variable-binding operator from the perspective of my way of tweaking the interpretation of the Kripkean semantics to reconcile it with actualism. Does it require us to do our semantics by choosing the (A), (QA) package over the (U), (QU) package? I think it provides strong motivation for opting for predication and quantification only over what exists.

One of the things that our theory of propositions brings out is the fact that our semantic and metaphysical theorizing is being done in a certain possible world (the actual world—the only one there is, if we really mean *world*) and that the semantics itself involves an ontological commitment to the objects to which the semantics appeals. Any theorist will agree with the obvious fact that we are doing our theorizing in the actual world, but the point is made salient by the assumption that many of the entities to which our semantics appeals—propositions, properties, functions, and relations, as well as individuals—are contingently existing entities. It is technically feasible to do the compositional semantics in a way that mixes reference to real things with artifacts of the model and then use the equivalence relations to filter out the artifacts at the end, but one might hope to do the semantics more directly, where all the values of constituents of a complex expression (be they individuals, properties, or higher-level functions from functions to functions . . . , or whatever) are entities that exist in the actual world. Our Tarskian semantics makes reference to infinite sequences of possible individuals, sequences that may include “merely possible individuals” that are artifacts of the model. But

this was just a notational convenience. Using the methods sketched in appendix B, we could do the compositional semantics directly, where all intermediate values in the composition are actual things. But this could not be done if our quantifier ranged over merely possible things. In any case, what the quantifiers in our regimented language are modeling is our commitment to the entities that are the ground for all that, according to our empirical and metaphysical beliefs, there is to truly say, and this is reason enough to restrict the range of our quantifier to what there is, in the widest sense.

Modal Realism, Modal Rationalism, Modal Naturalism

David Lewis promised us a paradise if we bought his modal realism, but most met the offer with an incredulous stare (which he regarded as a good, if not sufficient, reason to reject the offer). The offer might seem to have a Mephistophelian flavor: philosophical riches in exchange for your ontological conscience, your philosophical soul. Russell might have talked of the advantages of theft over honest toil, but let's not get too moralistic here. We know the advantages of theft over honest toil, but what is the downside? Who are the victims of our philosophical thievery? Is someone going to put us in jail for ontological profligacy or excessive postulation? Why should hard work make the goods we get any better? If modal realism is to be rejected, it must be (as is perhaps the case with Mephistopheles' offer) that the goods are not quite what they seem.

The riches Lewis promised (clarification and unification of a wide range of philosophical concepts) attest to the centrality of modal notions, and I am in full agreement with him that getting clear about the foundations of a theory of necessity and possibility can pay dividends across a wide range of philosophical issues in epistemology, metaphysics, philosophy of language and mind, and theory of action. All the more reason to get the foundations right, since if the promised riches are counterfeit, this may distort much else in our philosophy. Now I think it is clear that a lot of the constructive philosophical work that Lewis did in a modal realist

framework—work that he took to show the benefits of his metaphysical hypothesis—was immensely fruitful and clarifying, but I think the fruitful work can be separated from the metaphysical assumptions of that framework. Lewis (unlike Mephistopheles) was good at showing those who assessed the costs and benefits differently than he did how to buy some of what he had to offer without taking it all. But on some central applications of “modal realism at work,” I think the hypothesis of a plurality of worlds contributes to a misleading picture.

The modal realist offers accounts of *properties*, and this is a notion that I agree is centrally connected with modality. The framework I have been promoting takes the notion of a property as basic, and I argued that it is an essentially modal concept: one understands what a particular property is by understanding what it would be for it to be exemplified. Lewis, in contrast, gave a nominalistic account of properties that was independent of any modal notion. Properties are identified with their extensions. The universe, on this metaphysical picture, is nothing more than one big domain: reality consists just of individuals, and the collection of all of them can be divided up into sets in any way you like. Lewis’s modal realist picture does allow for distinctions between the sets that are the properties and between the sets of *n*-tuples that are the relations: some sets are natural or more natural than others. Some relations are spatio-temporal, while others are not. Using the notion of a spatio-temporal relation, one can divide the big domain into subdomains, equivalence classes of *world-mates*: things that are spatio-temporally related to each other. These subdomains are the possible worlds to which modal concepts are reduced. But being *natural* and being *spatio-temporal* are just higher-order properties and thus just sets of sets. They are sets of sets that have a special metaphysical status, one that is primitive and unexplained. Explanation has to stop somewhere, but it is not clear that this is a good place to stop. One may wonder what the ontological extravagance

required by Lewis's nominalism has bought us, if we end up swallowing the bitter ideological pill anyway.

The virtues of the modal realist's extensional metaphysics come from the clarity it brings to the questions of the identity conditions for properties, relations, propositions, and possibilities. This clarification motivates the project of modeling modal concepts with extensional models, but my aim in this book has been to ground those models in a theory that is not committed to the nominalistic metaphysics.

Another putative benefit of modal realism is that it offers an account of *propositional content*, and the account I am promoting agrees that modal notions are centrally connected with the idea of propositional content. My account agrees with Lewis's that propositions should be individuated by the way they distinguish between possibilities—by their truth conditions. There may be finer-grained notions of proposition that will play a role in the explanation of intentionality, but a notion of proposition with these identity conditions will be presupposed by any finer-grained notion. In the framework that I have been promoting, we started with an abstract space that was intended to model the structure of the propositions, individuated this way, that are the contents of our thought and speech. I emphasized in a number of places that propositions and possibilities are not representations—not things *with* truth conditions but truth conditions themselves. It is essential that they not be explained as representations if they are to play the right role in the explanation of intentionality. To clarify intentionality—the capacity to represent—we need to separate questions about what there is to be thought or said from questions about the capacities of the things in the world that are able to think or say those things. One thing we want a theory of intentionality to do is to distinguish a conception of the world as it is in itself from an account of the perspective of a thinker who has such a

conception,¹ and a representation of a possibility space is supposed to be a framework for doing this. But the modal realist picture, which adopts the same notion of content as a way of distinguishing between possibilities, blurs this line by treating distinctions between possible worlds like distinctions between places and thus treating all propositional content as indexical, perspectival, or self-locating content. On this theory, we can talk about where we are in reality, but there is nothing substantive that can be said about what reality is like in itself.²

But this book has not been a critique of modal realism; for the most part I have taken actualism for granted, aiming to develop a framework for clarifying semantic and metaphysical controversies that divide theorists who agree about actualism. There is, however, one component of Lewis's modal realism that is also a feature of many actualist accounts of modality, a doctrine that is often taken for granted. My main aim in the book has been to develop a framework that is compatible with the denial of this doctrine.

With some misgivings, I use the labels "modal rationalism" and "modal naturalism" for the contrast I have in mind. The former term is due to David Chalmers, and his thesis (or family of theses) is characterized in epistemological terms: the modal rationalist is one who identifies what is possible with what is conceivable; the central claim of the doctrine is that we have a priori access to modal truth. My concern is with a purely metaphysical thesis, but it is one that has epistemological consequences and that is presupposed by the doctrine that Chalmers characterizes, as well as by Lewis's modal realism: it is the thesis that all the distinctions between possibilities that could possibly be made are distinctions that exist from the perspective of any possible world—distinctions that exist necessarily. Modal naturalism is the contrasting doctrine,

¹ I discuss this issue more explicitly in Stalnaker 2008.

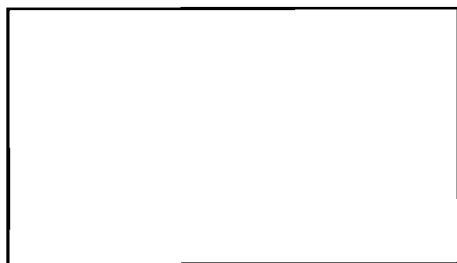
² This point is discussed in more detail in Stalnaker 2003, ch. 2.

though I suspect that some who agree with me about the doctrine may not agree that this is an appropriate name for it. Much of the discussion in this book has been about ways of avoiding modal naturalism. One strategy (Plantinga's) is to postulate primitive individual essences, another (that of Williamson and Linsky and Zalta) is to claim that what does exist and what might exist coincide. A third is to adopt an anti-haecceitist metaphysics that permits the reduction of individual thisness to qualitative character. My main aim has been to remove a motivation for modal rationalism, in all of these forms, by showing how to reconcile the orthodox modal semantics with the modal naturalist thesis and thus to have the virtues of an extensional semantics without the metaphysical baggage that it seems to bring with it. In most of the discussions of the modal rationalist strategies, the focus is on concrete particulars that seem to be contingent existents and on singular propositions about them, but these are only the easiest cases for motivating modal naturalism. As I have suggested, we need to recognize that qualities, relations, and some purely general propositions are also things that may exist only contingently, and so the framework I am proposing is a general abstract framework for modal semantics in which any contingent proposition might be a proposition that exists only contingently. The framework is not committed to modal naturalism. The metaphysical theories that result from implementing the various modal rationalist strategies can be stated within the framework and evaluated on their own terms. My aim was only to defuse certain arguments that purport to show that, on pain of incoherence, we must adopt some form of modal rationalism or give up on modal concepts altogether.

But modal naturalism has its own tensions and problems that are not fully resolved by the abstract framework. I will conclude by sketching a little more of the intuitive picture and then mentioning one of the problems that it reveals. Some may see the problems raised by modal naturalism as further reasons to try to avoid

it—to follow one of the strategies of populating the actual world with enough metaphysical exotica to ensure that all propositions exist necessarily. But I think the problems reveal real tensions in our modal concepts. Sometimes the problems that a theory brings to the surface should count as benefits, rather than costs, of the theory, though it will remain contentious when this is the right conclusion to draw.

Here is a picture of logical space:



The picture shows you everything that can be said, from an a priori point of view, about what is possible. In fact, the picture tells you far too much: it has a rich topological and metric structure, but all of that structure is just an artifact of the picture and not part of what is represented. Any permutation of the points in the space onto the points would be an equally good representation, and of course no significance should be attached to the fact that the picture is a rectangle. Furthermore, you should ignore the fact that the set of points has a certain cardinality—that is an artifact, too.

So our picture is an empty vessel—something to be divided up, in one way or another, to represent different ways that things might be. We can draw lines in the picture to represent real distinctions between a way things might be and the way they would be if things were not that way. But we cannot do this without using the resources available in the world we find ourselves in—a world that is represented, in our picture, by one of the points in the space. (But

of course nothing can be said, from an a priori point of view, about which of the points it is.)

So Adam finds himself in the presence of a large gray animal with tusks and a long trunk. "Let's call them 'elephants,'" he says, and he observes that the actual world is one in which there are elephants. Adam can then divide logical space into the part in which it is true that there are elephants and the part in which there are no things of this kind, and his observation that there are elephants is the observation that the actual state of the world is in the first part. If he had been on the other side of the line, in a world in which there are no elephants, he would not have been in a position to draw the line: he would have lacked the ability to make this distinction—it would have been inaccessible to him. But the modal naturalist goes further: perhaps the distinction Adam drew would not even exist if, contrary to fact, he had lived on the other side of the line. There is nothing terribly surprising about the claim that there are distinctions to be made that are inaccessible to some thinkers. For example, if we fictionally suppose, as Hilary Putnam did, that there is an actual but faraway planet, Twin Earth, that contains a substance superficially like water but radically different from water in its chemical structure, then we have supposed that there exists a substance that the inhabitants of earth are incapable of referring to. They can't think thoughts about it—can't draw the line between possible states of the universe in which this actual stuff is found somewhere and possible states in which it is not. But the incapacity of the earthlings in this fiction is just a human limitation. The universe as a whole (according to the fiction) contains the resources to ground the distinction: there are, in a possible world in which the fiction is true, facts of the matter about where in the universe this particular substance is found and whether it is found in the universe at all. But in the real actual world—the world in which Putnam's story is fiction—it is not only that there is no stuff, XYZ: there is no *kind* of stuff named by the pseudo-name XYZ that does

not exist but that might have existed. There is no property of being XYZ that the story identifies, which is not but might have been exemplified. This also is not so surprising. There are incomplete fictions that say there is a thing, or a kind, or a property meeting certain conditions but that fail to say enough to individuate any thing, kind, or property. But the modal naturalist holds that there are properties (like the property of being water, being an elephant, being Julius Caesar) that we succeed in individuating but that would not exist at all under certain counterfactual conditions. Perhaps, if there were no elephants, the proposition that there are no elephants would not exist.³

We have confronted, in the discussions in chapters 1 and 2, the intuitively paradoxical character of this conclusion, but I want to mention a more specific worry that the modal naturalist thesis raises. Start with the obvious fact that we are not omniscient. If we were omniscient, the activity of distinguishing between possibilities—of entertaining and expressing propositions—would have little point. But if what there is to be said or thought depends on which of the possibilities is actual, and if we are ignorant about that, then we may be ignorant about what there is to be said or thought. How can we know what we ourselves are thinking if the very existence of the content of our thought may depend on facts of which we are ignorant? Wittgenstein claimed, in the *Tractatus*, that “if the world had no substance, then whether a proposition had sense would depend on whether another proposition was true,” and “in that case, we could not sketch out any picture of the world (true or false).” The surrounding discussion in this part of

³I say “perhaps” to leave open the possibility that any particular property that fails to be exemplified might still be uniquely characterized in terms of things and properties that do exist. One might argue that the particular type of DNA that distinguishes elephants from other animal kinds could be specified, even in a world in which elephants did not happen to evolve, and one could argue that any animal with that DNA would be an elephant. But that is only to say that some properties can be analyzed in term of others.

the *Tractatus* makes clear that the “substance” that the world must have is a domain of necessarily existing objects that provides the “unalterable form” of the world.⁴ (So Wittgenstein opted for something like the Williamsonian version of modal rationalism.) Wittgenstein’s point, liberally interpreted, was that we can make sense of questions about where the world is in the space of possibilities only if we can form a conception of that space that is prior to and independent of the answers to the questions. This conclusion is, in my view, an overreaction to a genuine problem. It is right that to make sense of a representation of the world as being one way rather than another we need a standpoint that is neutral with respect to which of the ways it is. And it is right that if we opt for a version of modal naturalism, we must accept that there is no standpoint that is neutral with respect to all questions about contingent facts that might be raised. That is, there is no prospect of a division of the project of inquiry into two parts: first, a complete a priori articulation of the alternative possibilities, and second, an empirical investigation of questions about which of those possibilities is realized. Instead, we have to find ways of dividing up possibilities that are neutral with respect to what is at issue in a particular context, making factual presuppositions by restricting attention to a narrow range of alternative possibilities. Even if the propositions whose truth is at issue do not exist in all possible worlds, it will suffice if we can find propositions that exist in all of the possible worlds that are compatible with what is being presupposed. Adam can use reference to the elephants he sees to distinguish the possible worlds in which there are elephants from those in which there are not without knowing much about what elephants are like. That is good enough when the questions at issue are about how many there are, what they eat, whether they are dangerous. Things get more problematic when questions arise about the *nature* of elephants, since

⁴ Wittgenstein 1961, propositions 2.0211, 2.0212, and 2.023.

some possible worlds in which what Adam saw and named were creatures of a different kind are among the worlds on the other side of the line—worlds in which there are no elephants and in which the proposition that there are no elephants does not exist. It is not possible that elephants be reptiles, for example, but it may be possible that there be reptiles that look, to Adam, the way the elephants in fact looked to him, and it may be a metaphysical possibility that the kind that Adam named “elephant” was a reptile or, more plausibly, that it was a kind of mammal that differs in more subtle ways from the kind that Adam in fact named. To account for such epistemic possibilities, we need to fall back on other features of the actual world that are common to the relevant part of logical space, but our resources for distinguishing between the possibilities are rich and varied, and we can usually find ways of drawing the lines relative to facts that are reasonably presupposed in the particular context at hand. But what we can’t do, according to the modal naturalist, is to fall back on resources that are available from an *a priori* perspective. The devices we use to formulate questions about the nature of the things we find in the world should be seen not as a route to a defense of modal rationalism but as ways to help negotiate the difficult processes of inquiry and discussion when the identity and nature of things are at issue.

In these concluding remarks I am alluding to some ways that the contrast between modal naturalism and modal rationalism connects with many of the issues that have been the focus of attention in recent philosophical discussions about reference, intentionality, and knowledge: Frege puzzles about singular reference, the phenomenon of *a posteriori* necessities, the interpretation and application of two-dimensional modal semantics, contextualism about knowledge. This book has not been about those problems, but the development of the framework of modal semantics is motivated in large part by its relevance to them.

{ Appendix A }

Modeling Contingently Existing Propositions

Our aim is to construct a model for representing a space of propositions, and the potentialities that there be different propositions, all satisfying our modal theory of propositions. We do this with a space of points. One of the points is the designated actual world, and the actual propositions are represented by an algebra of subsets of the space (not the complete algebra of subsets but a subalgebra). Each point in the space is a potential “actual world,” and so each point will determine an algebra of subsets, which may be different for different points. Thus our model will have a family of equivalence relations on the space, one equivalence relation for each point. The different equivalence relations will be related to each other by certain structural constraints. The equivalence classes, or partition cells, defined by the equivalence relation for point x will be the maximally consistent propositions that would exist if x were the actual world.

The structural conditions that the family of equivalence relations must meet can be specified in either of two equivalent ways. I will describe the two formulations and then sketch the argument that they are equivalent.

Formulation I: W is the set of points, and for any x, y , and $z \in W$, $y \approx_x z$ says that y is equivalent to z , relative to x . Here are the structural conditions:

- (1) If $x \approx_x y$, then $y = x$. (This is the fixed point condition.)

(2) If $y \approx_x z$, then there exists a permutation function f from W onto W meeting these two constraints:

$$(2a) f(y) = z$$

(2b) for any u, v , and w , $u \approx_w v$ if and only if

$$f(u) \approx_{f(w)} f(v)$$

Condition 2, the structure-preserving condition, is necessary to ensure that equivalent points have the same representational significance. It does this by ensuring that the algebra of propositions that would exist if some point were realized is isomorphic to the algebra of propositions that would exist if some equivalent point were realized.

Formulation II begins not with a family of equivalence relations but with a family of permutation functions. In this formulation, it is assumed that for each $x \in W$, there is a set of permutation functions, F_x , mapping W onto W , that meets the following conditions. (I will represent the composition of two or more permutation functions by concatenation and the inverse of f by f^* .)

(1') If $f \in F_x$, then $f(x) = x$. (This is the fixed point condition.)

(2') F_x is closed under inverse and composition.

(3') If $f \in F_x$ and $g \in F_y$, then $fgf^* \in F_{f(y)}$. (This is the structure-preserving constraint.)

If we begin with formulation I, we can define the classes of permutation functions that are the primitives of formulation II as follows: F_x is the set of all permutation functions meeting these two conditions: (a) $f(y) = z$ only if $y \approx_x z$, and (b) for any u, v , and w , $u \approx_w v$ if and only if $f(u) \approx_{f(w)} f(v)$. One can then show that conditions (1'), (2'), and (3') are all satisfied for the defined sets of permutation functions. Condition (2b) is satisfied for any composition or inverse of permutation functions that satisfy it, and it is straightforward that the inverse and composition of permutation functions satisfying the fixed point condition for any given x will also satisfy

that condition. So since F_x is defined as all permutation functions meeting conditions (a) and (b), (2') follows. Here is the argument for (3'): As noted, (2b) is satisfied for any composition or inverse of functions that satisfy it, so the composition, fgf^* satisfies it. All that needs to be shown is that this composition satisfies the fixed point condition, relative to $f(y)$, given that (i) g satisfies it for y , and that (ii) f satisfies it for x . By (i) we have $g(y) = y$, and so substituting $f^*f(y)$ for y in $g(y)$, we have $gf^*f(y) = y$, from which it follows that $fgf^*f(y) = f(y)$, which is the fixed point condition we want.

If we begin with formulation II, we can define the relations $\approx_x$ for each x as follows: $y \approx_x z$ iff there exists an $f \in F_x$ such that $f(y) = z$. One can then show (using (2')) that the relations, defined this way, are equivalence relations and, using (3'), that the structure-preserving condition (2b) is satisfied.

The argument from (2') to the conclusion that the defined $\approx$ relations are equivalence relations is straightforward. The argument from (3') to (2b) is as follows:

Suppose for some x , y , and z , $y \approx_x z$. Then, by definition of ' $\approx$ ', $g(y) = z$, for some $g \in F_x$. Let f be any permutation function that is a member of F_w for some w . Then since $f^*fy = y$, we can conclude that $gf^*f(y) = z$, and so $fgf^*f(y) = f(z)$. Since by (3'), $fgf^* \in F_{f(x)}$, it follows from the definition of $\approx$ that $f(y) \approx_{f(x)} f(z)$.

{ Appendix B }

Propositional Functions and Properties

Properties, in standard models for a quantified modal language, are modeled by functions from possible worlds to subsets of the domain of the world; binary relations are (modeled by) functions from possible worlds to sets of ordered pairs of individuals from the domain of that world. Since the things we are modeling include properties and relations, and since a function is itself a special case of a relation, functions are also among the things that our models are modeling. A binary relation is a function iff for any x that is the first term of a pair in the extension of the relation (in a given possible world), there is a unique y that is the second term of a pair in the extension (in that world). A singular *propositional* function is just one kind of function, one with an individual as the argument and a proposition as its value. So it is a kind of binary relation and should be represented in the same way as binary relations between individuals. A singular propositional function, then, will be modeled by a function taking a possible world to a function taking an individual in the domain of that world to a proposition that exists in that world. If propositions themselves are modeled as functions (from possible worlds to truth values), then there will be three levels of function involved in the representation of a simple singular propositional function.

Singular propositional functions (which may be expressed in a quantified modal language by open sentences with one free variable) are thus very different objects from properties (which might

be expressed by a closed [simple or complex] one-place predicate). In a purely extensional semantics, the distinction between properties and propositional functions is a distinction without a difference: the value of a predicate is just a subset of the domain, and the value of a singular propositional function is a characteristic function taking a member of the domain to a truth value, and of course subsets of any domain correspond one-to-one to characteristic functions—functions from the domain to $\{1,0\}$. In the intensional case, however, the distinction matters. Let me illustrate the way the two notions can come apart with the distinction between the values of the predicates of existence and nonexistence on the one hand and the propositional functions expressed by open sentences asserting or denying existence on the other. The existence and nonexistence *predicates* can be defined in terms of identity with the help of a quantifier and a variable-binding predicate-forming operator as follows: $\hat{x}\exists\hat{y}x = y$ for existence and $\hat{x}\forall\hat{y}x \neq y$ for nonexistence. The values of these two closed one-place predicates are, respectively, the function taking each possible world to the whole domain of that world and the function taking each world to the empty set. It is a necessary truth that everything exemplifies the first of these properties and that nothing exemplifies the second. So the sentences $\Box\forall\hat{x}\exists\hat{y}x = y$ and $\Box\sim\exists\hat{x}\forall\hat{y}x \neq y$ are both valid. (Necessarily, everything exists, and nothing fails to exist.) But the propositional functions expressed by the open sentences $\exists\hat{y}x = y$ and $\forall\hat{y}x \neq y$ are functions whose extension, relative to a given world, is a function taking any member of the domain of that world to a proposition. For any world, and any member of the domain of the world, the proposition that is the value of the function will be a true proposition in the first case and a false proposition in the second case. But if the argument of the function (in a given possible world) is a *contingent* existent, then the propositions that are the values of the two functions will be, respectively, a *contingently* true and a *contingently* false proposition. It may be contingent that

an individual exists, but it is not thereby possible that it have the property of nonexistence.

One can apply a possibility operator to the open sentence, $\forall yx \neq y$, yielding an open sentence that expresses a singular propositional function that might in fact be satisfied. That is, the propositional function expressed by $\Diamond \forall yx \neq y$ might, for some arguments, take a value that is not only a contingent proposition but a true one. Then the closed predicate that results from binding the free variable in this open sentence ($\hat{x} \Diamond \forall yx \neq y$) will express a property that might be instantiated. Even though it is not possible that something have the property of nonexistence, a thing can have the property of possible nonexistence.

Despite the important difference between singular propositional functions and properties, there is still, in the intensional theory, a close connection between them. On the one hand, one can define a unique property in terms of any given propositional function: the property that a thing exemplifies (in a given possible world) if and only if the value of the given propositional function for that individual is a proposition that is true in that world. It is this correspondence that makes variable-binding work: complex (one-place) predicates are defined by an operator on open sentences, and the property expressed by $\hat{x}\phi$ is the property determined in this way by the propositional function expressed by the open sentence ϕ . On the other hand, one can also define a unique propositional function in terms of a given property: the function that takes the individual to the proposition that is true at a given possible world if and only if the individual has the property in that world. In the language, this fact is reflected in the fact that one can put a one-place predicate (expressing a property) together with a variable to get an open sentence that expresses the corresponding propositional function. But this correspondence between the two notions is not one-to-one. Different propositional functions may determine the same property.

The most basic variable-binding operator, and the only one that is needed, is an abstraction operator that operates on an open sentence (expressing a propositional function) and yields a (one-place) predicate, expressing the corresponding property. At least this is what happens in the case where the open sentence has just one free variable that is bound by the operator. In the general case, the predicate-forming operator yields a property function: a function taking an individual to a property or to a higher-level property function.

In an extensional theory, there is no need to distinguish a binary propositional function (taking a pair of individuals to a proposition) from a higher-order singular function, a function taking an individual to a singular propositional function. Each can be represented as a set of ordered triples, two individuals and a proposition. But in an intensional theory the difference matters: a function taking a single individual to a function taking a single individual to a proposition may differ from the corresponding function from a pair of individuals to a proposition. Since in the standard quantificational modal language variables are bound one at a time, and there may be modal operators between the variable-binding operators, we need the two-step procedure in defining the semantic values of complex expressions. One should, for example, think of an open sentence with two free variables as a singular function taking the value of one of its variables to a singular propositional function that takes the value of its other variable to a proposition. To illustrate the difference, here is an example which, like the one used earlier to bring out the difference between properties and propositional functions, involves the predicate of existence. The example is based on the following question George Bealer once put to me: how, without quantifying over possible individuals, can you express the obviously intelligible claim that there are possible individuals that are not compossible? (Here is an example, given by Nathan Salmon in a discussion of existence and nonexistence, of a

plausible claim of this kind. Consider a child that Nathan's parents might have had if the ovum from which Nathan actually developed had instead been fertilized by a different sperm cell. Nathan and this possible child [whom he dubs "Nothan"] are each possible, but the two are not compossible.)¹ My answer to Bealer's question was that what is said loosely in terms of possible individuals can be said more carefully as follows: There might have been an individual x such that there might have been an individual y such that it is impossible that both x and y exist. In symbols:

$$\Diamond \exists x \Diamond \exists y \Box \sim (Ex \wedge Ey)$$

(where E is the predicate of existence, defined as above). If one interprets the open sentence, $\Box \sim (Ex \wedge Ey)$, so that its value is a function taking a pair of individuals to a proposition, then its value, for any pair of individuals from the domain of any world, will be the impossible proposition. But if one interprets the open sentence as expressing a function taking an individual to a function taking an individual to a proposition, then when you bind the second variable with the existential quantifier, and then apply the possibility operator to it, you get a singular propositional function that may take an individual to a true proposition.

The two-step procedure generalizes in the obvious way: say that a level-0 propositional function is a proposition, and that a level $(k + 1)$ propositional function is a singular function from an individual to a level- k propositional function. Since variables are bound one at a time in standard quantified predicate languages, we need only singular functions of this kind.

If our variable-binding operator is a predicate-forming operator rather than a quantifier, then we need property functions as well as propositional functions. The predicate-forming operator will take

¹ Salmon 2005, 47.

level-1 propositional functions to properties but will also take level $(k + 1)$ propositional functions to level- k property functions.

It is clear enough, given the basic resources of a standard Kripke model, what level- k propositional functions and property functions are, and it is clear enough how the operators on propositions expressed by truth-functional and modal operators generalize to operators on level- k propositional functions, for all k . What requires further clarification, before we spell out the compositional semantic rules for the quantified modal language, is the following question: Which propositional and property functions, at which levels, are expressed by which open sentences and predicates? Suppose we have an open sentence with three free variables. It will express a propositional function of at least level 3, but which one? In a standard Tarskian formulation of quantification theory, the variables come with an order—a *global* order that is given in giving the vocabulary of the language, and the semantic rules for interpreting open sentences exploit this order. But this is not the order we want; the problem is that the order of the variables that is relevant to the interpretation of the open sentences is one that corresponds to the order in which variables are bound. So we do better with a *local*, context-dependent ordering of the variables, one that is defined for an open sentence, *relative to the complete closed sentence of which it is a part*. The procedure is straightforward. Say that a variable-binding operator has depth 1 in a given closed sentence if it does not occur in the scope of any other variable-binding operator and that it has depth $k + 1$ if it occurs in the immediate scope of an operator of depth k . Say that a variable has *rank* k in the context of a given complete closed sentence if the operator that binds it has *depth* k .

Say that a sentence has rank k if it is in the immediate scope of a variable-binding operator of depth k . Then rank- k sentences express level- k propositional functions, with the rank- k variable

marking the argument place of the function that is the semantic value of that open sentence.² Similarly, a rank- k open predicate expresses a k -level property function. The order of the variables in an open sentence or predicate is given by their rank. The idea is to determine the order of the variables by a top-down procedure but then to determine semantic values of the complex expressions by the usual bottom-up procedure. If you like, you can connect a global ordering of the variables (as in a Tarskian semantics) with this local ordering by relettering bound variables in the appropriate way for any sentence before applying the compositional rules to determine its value in a model.

Now for the compositional rules: The rules specify the propositions, properties, and propositional and property functions expressed by all the expressions, relative to an S5 Kripke structure³ and a valuation function. The valuation function takes the primitive descriptive expressions (predicates, names, and sentence letters) to functions from possible worlds to extensions. (The valuation function gives values to all predicates and sentence letters but may be a partial function with respect to names.) We will first state the rules for closed sentences and predicates, and then extend them inductively to give values to the rank- k open expressions for each k . The rules for atomic closed sentences and closed sentences with propositional operators and connectives are standard and straightforward. Here is the rule for complex predicates: If $\hat{x}\phi$ is a closed

²It may happen that the rank k variable does not occur in a sentence of rank k , since vacuous variable-binding is possible. For example, in the predicate $\hat{x}p$, the sentence letter p expresses a level-1 propositional function whose value, for any value of x , is the proposition expressed by p . Or in the context of the sentence $\forall \hat{x}\forall \hat{y}(Fx \vee Rxy)$, the open sentence Fx expresses a level-2 propositional function that takes any value of y to the level-1 propositional function that takes the value d of x to the proposition that d has the property expressed by F .

³I say "S5" because we are assuming that our models of metaphysical necessity are S5 models, with a universal accessibility relation, but the kind of semantics I am sketching works for any kind of Kripke structure.

predicate, then $v(\phi)$ is a level-1 propositional function, and so if d is an individual, $v(\phi)(d)$ is a proposition, and so $v_w(\phi)(d) \in \{1,0\}$. The extension for the predicate, relative to an arbitrary world w , is as follows:

$$v_w(\hat{x}\phi) = \{d \in D_w : v_w(\phi)(d) = 1\}$$

The quantifier, $\forall$, is not a variable-binding operator but an operator that operates on a rank- k predicate and yields a rank- k sentence. For closed sentences of the form $\forall\phi$ (where ϕ is a one-place closed predicate), the rule is as follows:

$$v_w(\forall\phi) = 1 \text{ iff } v_w(\phi) = D_w.$$

Finally, we will extend the rules to cover open predicates and sentences by first defining an extension of the language and the valuation function and then specifying the value of a level $(k+1)$ expression of the original language in terms of the value of a corresponding expression of the extended language of level k . Suppose that v is the original valuation function, that a is a new singular term, not in the language interpreted by v , and that d is any possible individual. $v^{(d/a)}$ extends the given v by specifying a value for the new term a :

$$v^{(d/a)}_w(a) = d \text{ (if } d \in D_w; \text{ it is undefined otherwise).}$$

For all primitive expressions ϕ of the original language, $v^{(d/a)}(\phi) = v(\phi)$.

Now we give a general rule for all open sentences or predicates of rank $k+1$ in terms of the values, relative to the extended valuation function, of expressions of the extended language of rank k . If we substitute a name for the $k+1$ level variable x in a sentence or

predicate of rank $k + 1$, the resulting expression will be a sentence or predicate of rank k (or less), so we can interpret it as expressing a level- k propositional or property function. Since we have already given values for all expressions of rank 0, this inductive clause will suffice for all the others. We specify a function by giving its value, for an arbitrary argument: For any expression ϕ of rank $k + 1$ and individual d , if x is the variable of rank $k + 1$ in ϕ , then

$$v_w(\phi)(d) = v^{(dia)}_w(\phi^a/x).$$

In the standard Tarskian semantics, the values of all sentences are taken to be infinitary propositional functions, and each of the compositional rules is given for all sentences, open and closed, rather than using the kind of inductive definition we have given. The semantic rules for sentential operators are defined from the start in terms of functions on infinitary propositional functions, and the semantic rules for variable-binding operators are stated in a way that uses just the global order of the variables. In effect, one treats the value of a sentence as something that determines values for all the different k -level functions that might be relevant to the interpretation of the sentence. The Tarskian procedure can be carried over to modal quantification theory, and in a standard Kripke model that uses this procedure, the arguments for the propositional functions are infinite sequences of *possible* individuals. By quantifying over all possible individuals, we can avoid distinguishing n -level propositional functions from n -ary propositional functions. But the use of infinite sequences, and of a global order for the variables, is just a technical convenience—it's just one way to do the bookkeeping. There is some conceptual distortion in this procedure, but it gets the job done. To preserve the comfort of familiarity, we will stay with the usual Tarskian notation when we discuss variable-binding and quantification in chapter 4. We can

talk with a clear conscience, in the metalanguage, about a domain of possible individuals because we have shown how to reconcile that talk with more austere ontological commitments and how to do the compositional semantics in a way that assigns as values only properties, relations, and functions that actually exist, according to the metaphysics that is presupposed.

{ Appendix C }

A Model for a Mighty Language

Suppose we have a standard S5 Kripke model, with a denumerable set of worlds, modeled by a denumerable set of maximal consistent sentences of our language. Being a mighty language, we assume that it has a name for each possible individual. In the general case, our modal language permits nonrigid singular terms, but in our mighty language, we will assume that all names are rigid so as to clearly separate the representation of the qualitative (by primitive predicates and expressions definable in terms of them in purely general terms) from the nonqualitative (represented by expressions that involve names). For convenience, I will also assume that each individual has only one name so that the names can model the possible individuals. Our aim is to specify a model for the mighty language that is not itself committed, either to haecceitism or its opposite, so as to state the anti-haecceitist thesis as one that distinguishes some mighty languages from others. I will define a special case of the kind of structure described in appendix A. We assume that all purely qualitative properties, relations, and propositions exist necessarily, so if there are contingently existing properties and propositions, they will involve "primitive thisness," to use Adams's term, and a rejection of the anti-haecceitist doctrine.

As is appropriate for a mighty language, we assume that our model is a *distinguished* model, in the sense that two worlds in it are distinct only if some sentence is true in one and not the other. So as with models constructed for a completeness proof, there will

be a one-to-one correspondence between the worlds and the sets of sentences true in them. Since this is a denumerable model, I will use the natural numbers to denote worlds and Γ_i , Γ_j , and so forth for the sets of sentences true in i , j , and so forth.

We will add to this model a family of permutation functions and then show that this family satisfies the conditions stated in appendix A. Start with the class of all permutation functions on names. Then extend these functions to permutations of sentences and sets of sentences: If f is a permutation function on names and ϕ is a sentence, let $f(\phi)$ be the result of substituting $f(a)$ for a , for every name a . Then for every set of sentences Γ , let $f(\Gamma) = \{f(\phi): \phi \in \Gamma\}$. Since worlds can be represented by sets of sentences, some of these permutation functions will determine functions mapping worlds onto worlds. Our family of permutation functions will be a subset of those.

Specifically, for some permutation functions f of the kind we have defined, it will be true that for every i , there will be a j such that $f(\Gamma_i) = \Gamma_j$. Call functions meeting this condition *closed* permutation functions.

Now for each world i we can define the appropriate class of permutation functions, F_i , that meet two conditions: $f \in F_i$ if and only both (i) f is closed, and (ii) $f(a) = a$ for all names $a \in D_i$. The classes of permutation functions, defined this way, will satisfy the structural constraints given in appendix A. This is easiest to see with formulation II: note that any inverse of a closed permutation function, or composition of two closed permutation functions, will be closed, and any inverse or composition of permutation functions satisfying condition (ii) will satisfy condition (ii), so our class F_i satisfies condition (2') (of formulation II, appendix A). We can show by an inductive argument that any closed permutation function satisfying condition (ii) will meet the fixed point constraint, which is the claim that for $f \in F_i$, $f(\Gamma_i) = \Gamma_i$. Note that $D_i = D_{f(i)}$. It

follows that, for every atomic sentence ϕ , $\phi \in \Gamma_i$ iff $\phi \in f(\Gamma_i)$: this holds for the case where $a \in D_i$ for all a in ϕ because then $\phi = f(\phi)$. But if for some a in ϕ , $a \notin D_i$, then also $a \notin D_{f(i)}$. So ϕ will be false in i and $f(i)$, and so not a member of either Γ_i or $f(\Gamma_i)$. The induction step for the truth-functional connectives is straightforward. In the case where ϕ has the form $\Box\psi$, then by S5 we have that $\phi \in \Gamma_i$ iff $\phi \in \Gamma_j$ for all i and j , so $\phi \in \Gamma_i$ iff $\phi \in f(\Gamma_i)$. For sentences of the form $\forall\phi$: given our requirement that every individual have a name, we know that for any j , $\forall\phi \in \Gamma_j$ iff $\{\phi^a/x : a \in D_j\} \subseteq \Gamma_j$, so $\forall\phi \in \Gamma_i$ iff $\forall\phi \in f(\Gamma_i)$. For sentences of the form $\hat{x}\phi(a)$: since any a is a rigid designator, we have $\hat{x}\phi(a) \in \Gamma_j$ iff $\phi^a/x \in \Gamma_j$, so by hypothesis of induction, $\hat{x}\phi(a) \in \Gamma_i$ iff $\hat{x}\phi(a) \in f(\Gamma_i)$. So the fixed point condition, $f(\Gamma_i) = \Gamma_i$ for all $f \in F_i$, is satisfied.

Condition (3'), the structure-preserving condition, is that if $f \in F_i$ and $g \in F_j$, then $fgf^* \in F_{f(j)}$. To show that this holds, it suffices to show that for any name a , if (i) $a = f(a)$ for all $a \in D_i$, and (ii) $g(a) = a$ for all $a \in D_j$, then (iii) $a = fgf^*(a)$ for all $a \in D_{f(j)}$. Suppose $a \in D_j$. Then by (ii), $g(a) = a$, and so $gf^*f(a) = a$. But then applying f to both sides we get $fgf^*f(a) = f(a)$ for all $a \in D_j$. Since $a \in D_j$ iff $f(a) \in D_{f(j)}$, we get (iii) $a = fgf^*(a)$ for all $a \in D_{f(j)}$.

For some models, the only permutation functions in the family will be the identity functions, and the corresponding equivalence relations will all be the identity relation. These will include the anti-haecceitist models in which the haecceities (properties expressed with a predicate of the form $\exists x x = a$) are all supervenient on the qualitative character of the world, but they will also include some models for which the supervenience thesis fails. For example, there are what we might call *Williamsonian* models, in which all possible worlds have the same domain. These might be haecceitist, since nothing excludes the possibility of two distinct but qualitatively indiscernible worlds in which distinct individuals that are actual in both worlds are interchanged. But we have also defined a more

generous set of permutation functions and a corresponding equivalence relation: the set of all *closed* permutations. Worlds that are equivalent in this sense are the worlds that are qualitatively indiscernible, so the anti-haecceitist models will be those for which this equivalence relation is identity.

Some models of our mighty language will be haecceitist, and some will not. Is the difference reflected in the logic? That is, are there consistent sets of sentences that are satisfiable only in haecceitist models? The answer is no, since nothing in the logic puts a limit on the predicates in the mighty language. We saw in chapter 3 that sentences such as

$$\exists \hat{x} \exists \hat{y} ((Fx \wedge Sy) \wedge \Diamond (Fy \wedge Sx))$$

were not satisfiable in an anti-haecceitist model *if our mighty language is such that $\exists \hat{x} \exists \hat{y} (Fx \wedge Sy)$ is maximal* (that is, if it determines a complete description of the world). But we can make that sentence, or any consistent set of sentences, true in an anti-haecceitist model by adding qualitative predicates to the language—by embedding it in a richer mighty language. This is easy to see, since we could add, for each haecceity $\exists \hat{x} x = a$, a new primitive one-place qualitative predicate A to the language, and if we add the sentence $\Box \forall \hat{x} (Ax \leftrightarrow x = a)$ to any consistent set of the original language, the result will still be consistent. Of course the predicates we add look suspiciously like those contrived Quinean predicates, “Socratizes” and “Pegasizes,” which Adams specifically excluded from his mighty language, but nothing in the abstract logic gives us a criterion for excluding such primitive predicates. The point underlines the fact that the mighty language method, whatever its heuristic value, does not give us an account of the suchness/thisness distinction but rather presupposes it. It also underlines the tenuousness of the distinction between a haecceitist, such as Plantinga, who holds that haecceities are primitive but exist independently of their

exemplifications, and an anti-haecceitist who holds that there are more qualitative properties than one might have thought. Suppose Plantinga were to categorize his individual essences as a special kind of qualitative properties. How would his metaphysical theory be different?

{ Appendix D }

Counterpart Semantics for the Cheap Haecceitist

Lewis preferred to formulate his counterpart theory as a first-order theory rather than a version of modal semantics, but a counterpart theorist might use a standard Kripke semantic framework, making just two changes to the structure and one to the semantic rules. The first change is to require that the domains of the possible worlds be pairwise disjoint. The second is to add to the model structure a counterpart relation, represented by a set of *counterpart functions*. A counterpart function takes a possible individual and a possible world to a possible individual from the domain of the world. (Counterpart functions may be partial.) Any binary counterpart relation between possible individuals can be represented by such a set of functions, but the function representation is more flexible, allowing for the kind of links between counterparts that Lewis wanted. (An example of the linkage: Suppose Hillary and Chelsea each have two counterparts in a given counterfactual world, *c*: H1 and H2 for Hillary, and C1 and C2 for Chelsea. C1 is the daughter of H1, and C2 is the daughter of H2 in *c*. Even though the pair $\langle H1, C2 \rangle$ is a pair of counterparts of the actual pair, it is not a counterpart of the pair. So there will be one counterpart function in the set that takes (Hillary, *c*) to H1 and (Chelsea, *c*) to C1 and another one that takes (Hillary, *c*) to H2 and (Chelsea, *c*) to C2 but no one function in the set that takes (Hillary, *c*) to H1 and (Chelsea, *c*) to C2.)

In Lewis's first statement of counterpart theory, it was required that each possible individual have exactly one counterpart—*itself*—in

its own possible world. In the counterpart function formulation, this requirement is that for any admissible counterpart function f , if $d \in D_w$, then $f(d, w) = d$. But as noted in the text, this requirement was later relaxed. Call a counterpart function *normal* if it satisfies the original requirement. The relaxed requirement is just that there exists at least one normal counterpart function in the set.

Following Lewis, we define a *possibility* as a pair consisting of a possible world and a counterpart function from the set. We can then define a single counterpart function f^* taking a possibility (rather than just a possible world) and a possible individual to a possible individual as follows: $f^*(d, \langle w, f \rangle) = f(d, w)$.

The semantics for our counterpart theory will be the standard Kripke semantics, with just two changes. First, possibilities will play the role of possible worlds. So, for example, in the Tarski formulation of the counterpart semantics, the semantic value of a sentence will be a function taking a sequence of possible individuals and a *possibility* (rather than a possible world) to a truth value. Second, we change the rule for assigning values to variables. Instead of the standard rule, $v_w^s(x) = s(x)$, we have $v_{\langle w, f \rangle}^s(x) = f(s(x), w)$.¹ (If $f(s(x), w)$ is defined; undefined otherwise.) All of the other semantic rules remain exactly the same, except with $\langle w, f \rangle$ replacing w . The quantified modal logic validated by the counterpart semantics is weaker than the logic for the classical theory, but the differences arise only when variable-binding operators and modality are mixed. The S5 propositional modal logic and the extensional logic of quantification and identity remain the same.²

¹ In the alternative semantics that is sketched in appendix B, the difference comes in the general inductive rule that defines the value of an expression of rank $k + 1$. This rule gives the value of a propositional or property function for an arbitrary argument, d , by extending the valuation function so that the value d is given to one new singular term. The rule for the regular Kripke semantics was this:

$$v_w^{(d/a)}(a) = d \text{ (if } d \in D_w; \text{ it is undefined otherwise).}$$

The corresponding rule for the counterpart semantics is this:

$$v_{\langle w, f \rangle}^{(d/a)}(a) = f(w, d) \text{ (if } f(w, d) \text{ is defined; it is undefined otherwise).}$$

² See Stalnaker 1994 for a discussion of a counterpart-function semantics.

Now we can add a one-place sentence operator, @, to our language, with the intended interpretation, "in an *actual* possibility." The semantic rule for it will be as follows:

$$v^s_{\langle w, f \rangle}(@\phi) = 1 \text{ iff } v^s_{\langle a, f \rangle}(\phi) = 1, \text{ for all normal counterpart functions, } f' \text{ (where } a \text{ is the actual world).}$$

The idea is that an *actual* possibility is a pair where the world is the actual world, and everything in the world is its own counterpart. Formally, this works perfectly well, but it changes the logic of the actuality operator in a way that shows it does not capture our intuitive notion. Because there may be more than one normal counterpart function, there may be distinct "actual" possibilities, by the definition we have given. So the problem brought out in Fara's argument is still there, even if we distinguish the actual world from the actual possibilities. Neither of the basic principles of the logic of actuality,

$$v^s_{\langle w, f \rangle}(\sim @\phi) = v^s_{\langle w, f \rangle}(@\sim \phi), \text{ or}$$

$$v^s_{\langle w, f \rangle}(@(\phi \wedge \psi)) = v^s_{\langle w, f \rangle}((@\phi \wedge @\psi)),$$

will be validated.

{ References }

- Adams, R. 1979. "Primitive Thisness and Primitive Identity." *Journal of Philosophy* 76, 5–26.
- . 1981. "Actualism and Thisness." *Synthese* 49, 3–41.
- Barwise, J., and J. Perry. 1985. "Shifting Situations and Shaken Attitudes." *Linguistics and Philosophy* 8, 105–61.
- Bennett, K. 2005. "Two Axes of Actualism." *Philosophical Review* 114, 297–326.
- . 2006. "Proxy 'Actualism.'" *Philosophical Studies* 129, 263–94.
- Black, M. 1952. "The Identity of Indiscernibles." *Mind* 61, 153–64.
- Carnap, R. 1950. "Empiricism, Semantics and Ontology." *Revue Internationale de Philosophie* 4, 208–28.
- Divers, J. 2002. *Possible Worlds*. London: Routledge.
- Fara, D. 2009. "Dear Haecceitism." *Erkenntnis* 70, 285–97.
- Fine, K. 2005. *Modality and Tense: Philosophical Papers*. Oxford: Oxford University Press.
- Grünbaum, A. 1964. "Is a Universal Nocturnal Expansion Falsifiable or Physically Vacuous?" *Philosophical Studies* 15, 71–79.
- Hacking, I. 1975. "The Identity of Indiscernibles." *Journal of Philosophy* 72, 249–56.
- Hawthorne, J., and T. Sider. 2002. "Locations." *Philosophical Topics* 30, 53–76.
- Hazen, A. 1979. "Counterpart-Theoretic Semantics for Modal Logic." *Journal of Philosophy* 76, 319–38.
- Hoffmann, A. 2010. "Response to Robert Stalnaker." In *Modality: Metaphysics, Logic and Epistemology*, ed. B. Hale and A. Hoffmann. Oxford: Oxford University Press, 33–34.
- Hofwebber, T. 2005. "Supervenience and Object Dependent Properties." *Journal of Philosophy* 102, 1–28.
- Kaplan, D. 1975. "How to Russell a Frege-Church." *Journal of Philosophy* 72, 716–29.
- King, J. 2007. *The Nature and Structure of Content*. Oxford: Oxford University Press.

- Kripke, S. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Lewis, D. 1986. *On the Plurality of Worlds*. Oxford: Basil Blackwell.
- . 1990. "Noneism or Allism?" *Mind* 99, 23–31.
- Linsky, B., and E. Zalta. 1994. "In Defense of the Simplest Quantified Modal Logic." *Philosophical Perspectives* 8, 431–58.
- Lowe, E. 1999. *The Possibility of Metaphysics: Substance, Identity and Time*. Oxford: Oxford University Press.
- McMichael, A. 1983. "A Problem for Actualism about Possible Worlds." *Philosophical Review* 92, 49–66.
- Peacocke, C. 2002. "Principles for Possibilia." *Royal Institute of Philosophy Supplement* 51, 119–45.
- Plantinga, A. 1983. "On Existentialism." *Philosophical Studies* 44:1–20. Reprinted in Plantinga 2003, 158–75.
- . 2003. *Essays in the Metaphysics of Modality*. Oxford: Oxford University Press.
- Priest, G. 2005. *Toward Nonbeing: The Semantics and Metaphysics of Intentionality*. Oxford: Oxford University Press.
- Prior, A., and K. Fine. 1977. *Worlds, Times and Selves*. London: Duckworth.
- Quine, W. 1948. "On What There Is." *Review of Metaphysics* 2, 21–38. Reprinted in Quine, *From a Logical Point of View*. Cambridge, MA: Harvard University Press, 1961.
- . 1953. "Three Grades of Modal Involvement." In *Proceedings of the XIth International Congress of Philosophy*. Amsterdam: North-Holland Publishing Company. Reprinted in Quine, *"Ways of Paradox" and Other Essays*. New York: Random House, 1966.
- . 1960. *Word and Object*. Cambridge, MA: MIT Press.
- Salmon, N. 1987. "Existence." *Philosophical Perspectives* 1, 49–108. Reprinted in Salmon 2005, 9–49.
- . 1998. "Nonexistence." *Noûs* 32, 277–319. Reprinted in Salmon 2005, 50–90.
- . 2005. *Metaphysics, Mathematics and Meaning*. Oxford: Oxford University Press.
- Sider, T. 2002. "The Ersatz Pluriverse." *Journal of Philosophy* 99, 279–315.
- Skow, B. 2008. "Haecceitism, Anti-Haecceitism and Possible Worlds." *Philosophical Quarterly* 58, 98–107.
- Stalnaker, R. 1976. "Possible Worlds." *Nous* 10, 65–65. Revised version reprinted in Stalnaker, 2003.

- Stalnaker, R. 1979. "Anti-Essentialism." *Midwest Studies in Philosophy* 4, 343–55. Reprinted in Stalnaker 2003.
- . 1986. "Possible Worlds and Situations." *Journal of Philosophical Logic* 15, 109–23.
- . 1994. "The Interaction of Modality with Quantification and Identity." *Modality, Morality and Belief: Essays in Honor of Ruth Barcan Marcus*, ed. W. Sinnott-Armstrong, D. Raffman, and N. Asher. Cambridge: Cambridge University Press. 12–28. Reprinted in Stalnaker 2003.
- . 2003. *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays*. Oxford: Oxford University Press.
- . 2008. *Our Knowledge of the Internal World*. Oxford: Oxford University Press.
- . 2009. "Merely Possible Propositions." In *Modality: Metaphysics, Logic and Epistemology*, ed. B. Hale and A. Hoffmann. Oxford: Oxford University Press.
- Van Fraassen, B. 1967. "Meaning Relations among Predicates." *Nous* 1, 481–95.
- Van Inwagen. 1998. "Modal Epistemology." *Philosophical Studies* 92, 67–84.
- Williamson, T. 2001. "Necessary Existents." In *Logic, Thought and Language*, ed. A. O'Hear. Cambridge: Cambridge University Press, 233–51.
- Wittgenstein, L. 1961. *Tractatus Logico-philosophicus*. London: Routledge and Kegan Paul.

{ Index }

- abstraction operator. *See* variable binding operator
- "actual," 4–5
- actualism, ix, 3–4, 8, 13, 43, 87. *See also*
 - models, actualist
 - and anti-haecceitism. *See* anti-haecceitistic strategy; haecceitism, and actualism; haecceitism and Plantinga
 - as ersatzism. *See* possible worlds, as representations
 - and iterated modality. *See* modality, iterated
 - and modal rationalism, 129. *See also* modal naturalism
 - and possible worlds semantics, 2, 20, 114–16, 124, 139–48
 - proxy, 115–24
 - and realism about possible worlds. *See* possible worlds, realism about
- actuality operator. *See* haecceitism, cheap
- Adams, Robert
 - on haecceitism. *See* haecceitism, Adams version
 - on the identity of indiscernibles. *See* identity of indiscernibles; haecceitism, Adams version
 - and Leibnizian phenomenalism, 76 (footnote)
- modal semantics, 96–102
 - on suchness/thisness distinction. *See* properties, qualitative vs. non-qualitative
 - on whether there is a total universe, 12 (footnote)
- aliens. *See* nonactual objects
- analysis (philosophical), 3, 4–5, 30
 - of modality. *See* modality, analysis of
- anti-haecceitism. *See* anti-haecceitist strategy, the; haecceitism
- anti-haecceitist strategy, the, 43, 52–53, 113, 130
- Barwise, John and Perry, John, 7, 41–42
- Bealer, George, 142
- Bennett, Karen, 114, 116–20
- Black, Max, 73–77
- bundle theory (of individuals), 70–71
- Carnap, Rudolf, 89–90, 120
- conceptual neutrality, 5, 8, 87–88, 90–92, 133–34
 - of present account of modality, 36, 42, 50, 88
- cost-benefit methodology, 3, 4, 5–6, 8, 127, 131
- counterpart theory. *See* haecceitism, and counterpart theory

direct reference, 93–96

Divers, John, 3–4

ersatzism, 9–10, 38

existentialism. *See* propositions, object-dependent; propositions, singular

extensionality. *See* possible worlds semantics, extensionality of

Fara, Delia, 81–82, 84–85, 156

Fine, Kit

on contingency of propositions, 17 (footnote)

on inner and outer truth. *See* truth of vs. truth in a world

and modalism, 30

on a puzzle about existence. *See* puzzle about existence

Grunbaum, Adolf. *See* “universal nocturnal expansion hypothesis”

Hacking, Ian, 74–76, 78

haecceities. *See* properties, identity

haecceitism, 52–88

and actualism, 58–59, 85–86, 113.

See also anti-haecceitist strategy; haecceitism, and Plantinga

Adams version of, 56–57, 97–98

cheap, 54, 81–86, 154–56

and counterpart theory, 55–56, 66, 67 (footnote), 71–72, 74, 78–86, 154–56

and the identity of indiscernibles.

See identity of indiscernibles

Kaplan version of 54–56. *See also* Kaplan, David

Lewis version of, 57–59

and the mighty language. *See* mighty language

and Plantinga, 44, 95–96, 118–19 (footnote)

and qualitative/non-qualitative

distinction. *See* properties, qualitative vs. non-qualitative

and rigid metaphysical determinism, 55–56. *See also* anti-haecceitistic strategy

Hoffmann, Aviv, 27 (footnote), 38

Hofwebber, Thomas, 101. *See also* direct reference

Humean supervenience, 72–73, 77–78, 87

identity of indiscernibles, 54, 71, 73–81

intentionality, 9, 53, 88, 128, 135

Kaplan, David, 59–60, 64, 66, 70, 78.

See also haecceitism, Kaplan version

King, Jeffrey, 11

Kripke, Saul, ix, 14, 15, 17, 18. *See also*

Kripke's dice; Kripke models; Kripke semantics

on a finest partition of logical space, 12 (footnote)

on reference, 96

seventh son, 17, 20–21, 28, 44

Kripke's dice, 18–19, 27–28, 32, 64–66, 70, 71–72, 85–86

Kripke models. *See* models

Kripke semantics. *See* possible worlds semantics

Leibniz, 68 (footnote), 76. *See also*

Leibnizian grounding principle

Leibnizian grounding principle, 67, 69, 76–77, 78

Lewis, David, ix, 14, 16

on causation. *See* Humean supervenience

and cost-benefit methodology, 5–7. *See also* cost-benefit methodology

- on counterpart theory. *See* haecceitism, and counterpart theory
- and direct reference, 96. *See also* direct reference
- and ersatzism. *See* ersatzism
- on haecceitism. *See* haecceitism, cheap; haecceitism, Lewis version; haecceitism, and rigid metaphysical determinism
- and the identity of indiscernibles, 58–59, 71, 73, 78–81. *See also* identity of indiscernibles
- on the mighty language, 61. *See also* mighty language
- on modal realism. *See* modal realism
- and modalism, 30. *See also* modality, analysis of
- on properties, 6–7, 127–28. *See also* bundle theory (of individuals); mighty language
- on proposition individuation, 127
- Linsky, Bernard, and Zalta, Edward, 116, 119, 130
- logical space, 23, 26–27, 29, 31–32, 38, 41, 113–14. *See also* possible worlds
- and cheap haecceitism, 85–86
- finest partition of, 13, 25–26, 26–27. *See also* possible worlds semantics, intended model of
- identified with the property space, 70
- and the mighty language, 62
- and modal naturalism, 131–35
- Lowe, E. J., 22
- maximality. *See* possible worlds, maximality of; propositions, maximal
- McMichael, Alan
 - on extensionality. *See* possible worlds semantics, extensionality of
- on iterated modality. *See* modality, iterated
- on realism about possible worlds. *See* possible worlds semantics, realism about
- merely possible individuals. *See* non-actual objects
- “metaphysically satisfying, though formally inconvenient semantics.” *See* Adams, Robert, modal semantics
- metaphysics, x, 51, 53, 126
 - actualist. *See* actualism
 - and counterpart theory. *See* haecceitism, and counterpart theory
 - extravagant, 2, 44, 50, 118, 120
 - haecceitistic. *See* haecceitism
 - of intrinsic and relational properties. *See* Leibnizian grounding principle
 - of nonactual objects. *See* nonactual objects
 - possibilist. *See* modal realism
 - of possible worlds. *See* possible worlds
 - vs. semantics, 4–5, 53, 55–56, 59, 87–88, 89–125
- mighty language, 61–62, 62–66, 149–53
- modal naturalism, 129–34
- modal rationalism. *See* modal naturalism; modal realism, and modal rationalism
- modal realism, 10, 85, 123, 126–29
 - and “actual,” 5
 - and anti-haecceitism. *See* haecceitism, Lewis version
 - and cost benefit methodology, 6–7
 - and “the language of boxes and diamonds,” 83
 - and modal rationalism, 87, 129
- modalism. *See* modality, analysis of

modality, 2, 5–6

Adams on. *See* Adams, Robert,
modal semantics

analysis of, 3–4, 30, 49–50

iterated, 19–21, 31, 33, 35, 64

and modal rationalism. *See* modal
naturalism

and possible worlds. *See* possible
worlds

and properties. *See* properties

primitive, 30, 31

and variable-binding operators, 155.

See also variable binding operator

worldly vs. transcendent. *See* worldly
and transcendent necessities

modal semantics. *See* possible worlds
semantics

models (model-theoretic), 41. *See also*
Kripke's dice

actualist, 40, 41, 42, 62, 86, 113–14,
147–48. *See also* actualism

artifacts of, 2, 35, 37, 39, 64 (foot-
note), 124. *See also* models, of
logical space; models, of physical
space

extensional, 128

haecceitistic (or anti-haecceitistic),
63–64, 66, 83–84, 85, 153. *See also*
haecceitism

intended, 7, 40, 41–42

of logical space, 23, 29, 31–32, 32–33,
62, 128, 136. *See also* logical space;
models, of properties

of a mighty language, 149–53. *See*
also mighty language

of physical space, 33–34, 41, 75–76.
See also space (physical)

of properties, 35, 37, 62, 67–73, 139,
144–45. *See also* properties

of propositions. *See* models, of logi-
cal space

as representations, 38. *See also*
representation
used instrumentally, 7, 40–41

neutralism. *See* quantification,
non-committal

nominalism. *See* Lewis, David, on
properties

nonactual objects, 3, 114–15, 124,
147–48. *See also* compossibility
sentence; Kripke, Saul, seventh
son; possible worlds semantics,
domain of quantifier in

and actualism. *See* actualism, and
possible worlds semantics; haec-
ceitism, and actualism

and possible worlds semantics. *See*
actualism, and possible worlds
semantics

so-called, 21, 33, 39

nonactual possibles. *See* nonactual
objects

object dependence. *See* propositions,
object-dependent

Peacocke, Christopher, 2

Plantinga, Alvin

on individual essences, 2 (foot-
note), 50, 52, 130, 152–53. *See*
also haecceitism; possible worlds
semantics, and Plantinga

on object dependence, 42, 42–47,
93, 94, 95. *See also* propositions,
object-dependent

and possible worlds semantics. *See*
possible worlds semantics, and
Plantinga; properties, and predi-
cates on Plantinga semantics

Pollock, John, 45, 47

possibilism. *See* modal realism

- possible worlds, ix, 2, 3, 8, 18. *See also*
 logical space; possible worlds se-
 mantics; properties; propositions
 and analysis of modality. *See* modal-
 ity, analysis of
 contingency of, 3, 21, 22–23, 38, 47
 and haecceitism. *See* haecceitism;
 possible worlds, indiscernible;
 possible worlds, and property
 spaces
 indiscernible, 58–59, 65, 69, 152
 intrinsic vs. relational properties of,
 72–73, 87
 maximality of, 12–13, 19–20, 48, 85
 metaphysical innocence of, 2–3
 and modal realism. *See* modal
 realism
 as points vs. cells of logical space,
 13, 25
 as properties, 3, 8, 11, 12, 17, 37
 and property spaces, 69, 70
 and propositions. *See* propositions,
 as sets of possible worlds
 realism about, 7, 23, 32, 39. *See also*
 possible worlds semantics, realism
 about
 as representations, 9–11, 37, 58–59,
 66
 as sets of propositions, 49
 truth in vs. truth at. *See* truth in vs.
 truth at a world
- possible worlds semantics, ix, 29
 Adams version. *See* Adams, Robert,
 modal semantics
 and actualism. *See* actualism, and
 possible worlds semantics
 and analysis of modality. *See* modal-
 ity, analysis of
 and contingent propositions, 23, 130
 domain of quantifier in, 114–25
 extensionality of, 21, 23, 29–30, 130
- intended model of, 7
 and nonactual objects. *See* actualism,
 and possible worlds semantics;
 possible worlds semantics, do-
 main of quantifier in
 and Plantinga, 116, 117–18, 119–20,
 123
 realism about, 2–3, 7–8, 13, 21, 23,
 32–33, 35, 36, 39–42. *See also* pos-
 sible worlds, realism about
- Powers, Larry, ix, 45, 46
- predicates vs. operators, 99, 100 (foot-
 note), 105–6, 108, 139–47. *See also*
 propositional functions
- presentism, 122, 123–24
- Priest, Graham, 122
- Prior, Arthur, 28, 43, 44, 101, 101–3
- properties, 8, 11, 37, 42. *See also* Leib-
 nizian grounding principle; Lewis,
 David, on properties; possible
 worlds; propositions
 basic, 105, 123
 as constituents of propositions. *See*
 propositions, Russellian
 containment, 14–17, 28–29
 contingency of, 3, 133. *See also*
 possible worlds, contingency of;
 propositions, contingency of
 higher-order, 17, 18–19, 36–37, 41
 identity, 44, 56–57, 69, 113, 118, 151.
See also Plantinga, Arthur, on
 individual essences
- intrinsic vs. relational. *See* Leibniz-
 ian grounding principle; identity
 of indiscernibles
- Lewis on. *See* Lewis, David, on
 properties
- maximal. *See* possible worlds, maxi-
 mality of
- modal, 55–56, 60, 62, 65–66, 70,
 71–72, 85, 87

- properties (*continued*)
 and modal realism, 127–128. *See also* modal realism
 modeled by abstract spaces, 35, 67–71
 and predicates on Plantinga semantics, 117–118
 and propositional functions, 139–48, 155(footnote)
 qualitative vs. non-qualitative, 20 (footnote), 52–54, 56–66, 69–70, 87, 149, 152–53. *See also* anti-haecceitist strategy, identity of indiscernibles
 vs. representations, 3, 9–11, 37–38
 second-order. *See* properties, higher order
 spatial. *See* space (physical)
 specific, 15–17, 19
 propositional functions, 37, 95, 107, 139–48. *See also* predicates vs. operators
 propositions. *See also* logical space; possible worlds; properties
 complete domain of. *See* logical space, finest partition of
 contingency of, 3, 17 (footnote), 18 (footnote), 21, 22–23, 27, 38, 42–51, 52–53, 63–64, 86, 124, 130–33, 136–37. *See also* possible worlds, contingency of; properties, contingency of; propositions, modal properties of
 existential, 36, 53
 general. *See* propositions, qualitative
 higher order, 28–29, 31, 32
 maximal, 19–20, 27–28, 31–32, 32–33, 62, 87. *See also* possible worlds, maximality of
 merely possible, 28, 62
 minimal theory of, 23, 23–27, 29, 101
 modal properties of, 23, 27 (footnote), 29, 36. *See also* propositions, contingency of
 object-dependent, 22, 27 (footnote), 93–94. *See also* propositions, singular
 and operators, 99–100
 qualitative, 43, 52–53, 61, 63–64, 130, 149
 as representations, 10–11, 128. *See also* properties, vs. representations
 Russellian, 26, 43
 as sets of possible worlds, 12, 27, 43
 singular, 20, 22, 27 (footnote), 28, 42–43, 46, 52–53, 94, 95, 97, 130
 structured, 94, 96
 true in vs. true at a world. *See* truth in vs. truth at a world
 as truth-conditions, 11, 43, 128
 Putnam, Hilary, 13 (footnote), 96, 132
 puzzle about existence, 92, 102–13
 quantification, 4, 7, 30, 91–92, 106–7, 110
 and modal logic, 63, 93, 98, 113–25, 139–48, 155
 non-committal, 120–23
 Quine, W. V.
 on ontology, 1, 8, 120–21, 123
 on properties, 6
 on regimentation, 90–92. *See also* regimentation
 on three grades of modal involvement, 99, 100 (footnote)
 reduction. *See* analysis, philosophical
 regimentation, 90–92, 98, 102, 18, 109–10
 of modal discourse, 11, 40, 83, 92, 109–11, 125
 of quantificational discourse, 110–11, 125. *See also* quantification

representation, 9, 78, 90, 98, 128, 134
 and possible worlds. *See* possible
 worlds, as representations
 and properties. *See* properties, vs.
 representations
 and propositions. *See* propositions,
 as representations

Salmon, Nathan, 120, 121–23, 142

semantics

vs. metaphysics. *See* metaphysics, vs.
 semantics

possible world. *See* possible world
 semantics

sempiternal vs. eternal truths, 112

Skow, Brad, 9–10

space (physical), 53, 87

absolute, 68, 74, 76

idealism about, 76 (footnote)

modeled by abstract spaces, 33, 35,
 67

relational theory of, 33, 41, 68 (foot-
 note), 69, 75–76, 114

“spatial dispersal argument.” *See* iden-
 tity of indiscernibles

thisness vs. suchness. *See* properties,
 qualitative vs. non-qualitative

transcendent vs. worldly necessities,
 102, 111–13

truth, inner vs. outer. *See* truth of vs.
 truth in a world

truth in vs. truth at a world, 28, 46,
 47–48, 50, 101

Twin Earth, 132

“universal nocturnal expansion hy-
 pothesis,” 34

utility theory, 35

variable binding operator, 46, 107, 115,
 118–19, 124, 140–47, 155

Williamson, Timothy, 2, 45, 48–50, 51,
 116, 130

Wittgenstein, Ludwig, 1, 14, 15, 133–34

necessary existence of everything that does or might exist. In contrast, *Mere Possibilities* shows how we can make sense of ordinary beliefs about what might and must exist without making counterintuitive metaphysical commitments. The book also sheds new light on the nature of metaphysical theorizing by exploring the interaction of semantic and metaphysical issues, the connections between different metaphysical issues, and the nature of ontological commitment.

Robert Stalnaker is the Laurance S. Rockefeller Professor of Philosophy at Massachusetts Institute of Technology. He is the author of *Our Knowledge of the Internal World*, *Ways a World Might Be*, *Context and Content*, and *Inquiry*.

CARL G. HEMPEL LECTURE SERIES

Jacket design by Leslie Flis

Jacket photo: *High-resolution texture of red brush strokes on white background.* © GoodMood Photo. Courtesy of Shutterstock.

To receive emails in your area of interest, sign up at press.princeton.edu

Mere Possibilities



METAPHYSICAL FOUNDATIONS OF MODAL SEMANTICS

Robert Stalnaker

"This is a subtle, nuanced exploration of the metaphysical basis for talk of possible worlds and its far from straightforward relation to the version of formal semantics known as 'possible worlds semantics,' both of which have been immensely influential in recent decades. Robert Stalnaker has been one of the major players in debates on these matters, and this book contains significant further developments of his ideas."

—TIMOTHY WILLIAMSON, *University of Oxford*

"This is a first-rate book and a model of the best kind of unfussy scholarly writing. Robert Stalnaker combines an admirable economy of presentation with clarity, rigor, and thoroughness in exposition and argument. This is a major contribution to the field, and one to which philosophers, logicians, and linguists with any serious interest in modality should attend."

—BOB HALE, *University of Sheffield*



PRINCETON
press.princeton.edu

ISBN 978-0-691-14712-3



9 780691 147123