INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning 300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA



*

. . .

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

AN ESSENTIALIST THEORY OF MODALITY

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

RANDALL L. RIDENOUR Norman, Oklahoma 2000 UMI Number: 9962959

UMI®

UMI Microform 9962959

Copyright 2000 by Bell & Howell Information and Learning Company. All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

> Bell & Howell Information and Learning Company 300 North Zeeb Road P.O. Box 1346 Ann Arbor, MI 48106-1346

© Copyright by RANDALL L. RIDENOUR 2000 All Rights Reserved AN ESSENTIALIST THEORY OF MODALITY

-

A dissertation APPROVED FOR THE DEPARTMENT OF PHILOSOPHY

BY

n Reinaldo Elugarda James Hauthome

•

ACKNOWLEDGEMENTS

I owe a great debt of gratitude to many people who have made this possible. First, I would like to thank the members of my committee, Hugh Benson, Jim Hawthorne, Ray Elugardo, and Ralph Doty, for their invaluable contributions. I especially owe great thanks to Chris Swoyer, who directed this dissertation. His comments were, as always, both insightful and encouraging. I hope someday to gain his remarkable ability to make very difficult concepts clear and comprehensible.

I must also thank the faculty and graduate students at the Department of Philosophy of the University of Oklahoma for the stimulation and friendship that they have provided. Given my lack of administrative talent, I could never have completed my studies without the help of Susan Nostrand and Shelley Konieczny. I am grateful for all that they have done.

I would also like to thank my students and colleagues at Oklahoma Baptist University. They have, sometimes unknowingly, been sounding-boards for many of the ideas that are found here.

Finally, I owe everything to my wife, Sheri, and daughter. Rachael. Without their support and encouragement none of this would have been possible, nor would it have even been worthwhile. Ideas are only valuable insofar as they are shared with those you love.

CONTENTS

<u>l</u> .	THE PROBLEM OF NECESSITY	1
	Introduction	1
	The Problem	1
	The Solution	5
<u>2.</u>	MODALITIES: FROM LOGIC TO METAPHYSICS	10
	Introduction	10
	The Ancient Period	11
	The Medieval Period	23
	The Modern Period	36
	The Formal Systems of C.I. Lewis	45
	Quine's Modal Skepticism	52
	Kripke and Modal Semantics	
<u>3.</u>	POSSIBLE WORLDS: COMPETING PERSPECTIVES	65
	Introduction.	65
	The Case for Possible Worlds	66
	Modal Realism	76
	Modal Actualism	83
	Evaluating the Two Accounts: Realism vs. Actualism	98
<u>4.</u>	PROBLEMS FOR POSSIBLE WORLDS	114
	Introduction	114

.

	The Problem of Knowledge	116
	The Problem of Relevance	121
	The Problem of Coherence	
	The Paradox of the Liar	
	The Cantorian Argument	145
	<u>Conclusion</u>	149
<u>5.</u>	THE ESSENTIALIST SOLUTION	151
	Introduction	151
	Aristotelian Essentialism	
	Contemporary Modal Essentialism	164
	Evaluating Modal Essentialism	
	Definition and Essence	
	Essence and Modality	
	Individual Essences	
	Essences of Artifacts	195
	Knowledge of Essences	198
<u>6.</u>	OBJECTIONS AND REPLIES	
	Introduction	
	Natural Kinds	
	Relativism and Realism	
<u>CC</u>	DNCLUSION	
<u>Lľ</u>	TERATURE CITED	

CHAPTER 1

THE PROBLEM OF NECESSITY

Introduction

For several months following the 1996 United States presidential election, a friend regularly said to me, "You know that Bob Dole *could* have won that election." Although this statement seems true, it raises two particularly difficult philosophical questions. First, what feature of the world makes it true? Given that Dole lost, what makes it true that he could have won? Second, even if it is the case that Dole could have won, how can we ever come to know that it is true? The goal of this dissertation is to show that, although current approaches fail, these questions can ultimately be answered.

The Problem

As I survey the beliefs that I have, I find that some of them are quite ordinary. For instance, two of my beliefs can be expressed using the following English sentences:

(1) George Washington was the first president of the United States.

(2) A triangle has three sides.

These beliefs have little in common, but one important feature that they do share is that they are both true. Intuitively, they are true because they both express a fact about the world. George Washington was, in fact, the first President; and triangles do, in fact, have three sides.

As I continue the survey of my beliefs, I find that I have others, similar to these, but with an important difference. These beliefs are expressed by the following sentences:

(1*) George Washington *might not have been* the first president of the United States.

(2*) A triangle *must* have three sides.

Most would grant that these beliefs are also true. Even though the events of the late 18th century led to Washington's election, those events might have been different. Washington could have died in childhood, been killed in battle, or simply fallen into the Delaware and died of hypothermia. If any of these had occurred, then Washington would not have been the first President.

By contrast, it is not possible for a triangle to lack three sides. since triangles are defined as three-sided polygons. So, if there were a triangle with more than three sides, then there would be a three-sided polygon that did not have three sides. Intuitively, this contradiction shows that triangles must have three sides.

Now, consider the following two sentences, which are both false:

(3) The 1996 United States presidential election was won by a Republican.

(4) Some triangles have four sides.

These two false sentences are similar to the two true sentences which follow:

(3*) The 1996 presidential election *might* have been won by a Republican.

(4*) No triangle can possibly have four sides.

2

The truth of (4^*) is shown by the same argument for (2^*) . The argument for (3^*) is mostly an appeal to intuition, much like the argument for (1^*) . There were times during Bill Clinton's first term when many believed that he would never win reelection. Given the structure of our political system, if the Democrat would have lost, then, in all probability, the Republican would have won. Had Dole's campaign been run differently, then he quite possibly would have won the election.

The truth of the starred sentences plays a significant role in our attitude toward the original sentences. It is true that George Washington was the first President of the United States, but it could have been false. It is false that the 1996 United States Presidential election was won by a Republican, but it could have been true. It is true that a triangle has three sides, and it could not possibly be false. Finally, it is false that some triangles have four sides, and it could not possibly be true.

There is therefore more to a sentence than simply being true or false. Its truth can be a matter of necessity as in (2), or impossibility as in (4). A sentence could also be contingent, as in both (1) and (3), that is, it is possible that the sentence be true, but also possible that it be false. This feature of a sentence, that is, its *necessity*. *impossibility*. or *contingency*, is called its *modal status*.

Intuitions about necessity and possibility play an important role in our everyday practical reasoning. When planning a trip, one might consider if it is possible to drive from Oklahoma City to Dallas and back in one day and have time to see both the zoo and the museum of fine art. Before repairing a broken power cord on an appliance, one might wonder if it is possible that the appliance have an electrical current running through it even though the appliance is unplugged.

Modalities also play a crucial role in all areas of philosophy. An obvious example is the investigation of the logical relationships between sentences such as (2) and (2*) above. An even more fundamental concern is the notion of validity in logic. A valid argument is commonly defined as one in which it is impossible for the premises to be true and the conclusion false. In metaphysics, one might ask what reality must be like given the experiences that we have. The epistemologist, when addressing skepticism, is concerned with the possibility of knowledge. In ethics, if 'ought' implies 'can,' then to say that one ought to have acted differently is to say that one could have done so. If this is the case, then it seems that moral responsibility requires the possibility of doing other than what one in fact did. To deny this is just to claim that it is possible to be morally responsible and one's actions still be determined. In social theory, one might ask if it is possible to develop a just society if everyone in society is an egoist, only concerned with self-interest. An important issue in aesthetics is the possibility that a work of art is beautiful even though no one considers it to be so. These examples show the importance of the concepts of possibility and necessity in all areas of philosophical theorizing.

Given the importance of modal concepts to philosophy, one should immediately recognize the usefulness of a theory of modality in helping to systematize and explain our intuitions about possibility and necessity. It has only been recently, however, that such theories have been developed.

4

This is not surprising, given the empiricist leanings of contemporary philosophy. I can know that a democrat won in 1996, since I experienced a democrat winning in 1996. How can I know that a republican might have won in 1996, however, since all I experienced was a democrat winning?

This problem is an epistemological one, concerning how we can come to know the truth of sentences containing modal notions. An even more serious problem, however, is metaphysical. What feature of the world makes such sentences true? What makes the sentence 'A democrat won in 1996' true is that a democrat actually won in 1996. It is difficult, however, to point to any feature of the world that makes the sentence 'A republican might have won in 1996' true. These two problems, the metaphysical and epistemological, together comprise what Michael Dummett has called "the philosophical problem of necessity."

The Solution

In this dissertation, I will offer a solution to the problem of necessity that is both new and ancient. It is new in that it is radically different from most approaches to modalities in contemporary philosophy. It is ancient in that it urges a return to an approach advocated by Aristotle in the 4th century BCE.

In chapter two, I will survey the history of modal logic, and show how it has greatly impacted current theories of modality.¹ Aristotle was the first to develop a

¹ To enhance readability. I will adopt certain conventions of symbolization throughout this dissertation. I will omit unnecessary subscripts, superscripts, and parentheses. I will use the lower case Roman italic letters, p^{*} and q^{*} for metalinguistic propositional variables. I will abbreviate object language individual terms by lower case Roman letters, predicate terms by capital letters, using letters near the beginning of the alphabet for constants and letters near the end for variables. I will

systematic modal logic, and his work continued to be refined and improved throughout the medieval period. Our basic understanding of the logical relationships between modal sentences has changed little since the end of the medieval period.

Most current attempts to solve the metaphysical problem of necessity, however, can trace their intellectual lineage back to Leibniz. He analyzed the concept of necessity as truth in all possible circumstances or, as he picturesquely put it, truth in all possible worlds. Saul Kripke and others have since used this notion of necessity as truth in all possible worlds to develop a semantics for modal logic. As a heuristic, the concept of possible worlds has proven extremely useful in analyzing the meaning of statements using modal terms. Given this success, it is natural to think that statements about modalities are really about possible worlds. In other words, what makes the statement 'A triangle must have three sides' true is that there is no possible world in which a triangle does not have three sides. This is to say that modal facts are grounded in facts about possible worlds. An analysis of modalities, therefore, requires an analysis of possible worlds.

Chapter three is an examination of the two primary positions on possible worlds. One position, associated primarily with David Lewis but also advocated by others, is a realism about possible worlds. The realist, as Lewis defines it, believes that possible worlds exist, and are the same kind of thing as the actual world. Other

also adopt the convention of autonymous use, letting each simple expression stand for itself and each juxtaposition of expressions stand for their concatenation. Finally, names of particular logic systems are bold-faced, the names of their characteristic formulas are not (e.g. S-S refers to the system itself, while S5 refers to its characteristic formula).

worlds are like our world, they have people, animals, houses, and trees. The significant difference between our world and other worlds is perspectival: we are in this one, and not in the others. The others are no less real, however. It is tempting to immediately dismiss such realism as too far-fetched to be true. Realism. however, has an apparent strength that is difficult to dismiss. If realism were true, then it appears possible to explain modalities without using any language about possibility and necessity. This gives the realist theory a great deal of explanatory power.

The other type of position is actualism, of which there are many varieties. The actualist believes that possible worlds are constructed of actually existing objects, such as obtainable states of affairs or instantiable properties. An important point is that the actualist can explain possibility only in terms of a state of affairs that possibly obtains, or a property that is possibly instantiated. Actualism explicitly lacks much of the explanatory power of realism, because it is committed to a modal concept, either instantiability or obtainability, that is primitive and cannot be explained.

Possible world theories have proven to be very powerful tools. quite useful in addressing several previously intractable philosophical problems. These problems include not only the nature of modality, but also the nature of counterfactuals, intensional states, and properties. Unfortunately, these successes come at an extremely high price. Just as one is able to derive anything from a contradiction, one is able to achieve great philosophical success when one begins with a paradox.

In the fourth chapter, I will show how such theories, despite their great explanatory power, cannot succeed. There, I will argue that possible world theories suffer from three primary problems. First, they fail to provide a plausible account of how we come to know modal facts. Second, it is not clear how possible worlds as usually construed are even relevant to actual modalities. Third, and most importantly, all versions of possible world metaphysics are committed to a maximal set of true propositions, and such a set is incoherent.

In chapter five, I will argue that the only successful ground for modal facts is in the properties of actually existing objects, specifically their essential and nonessential properties. If this is right, then the belief that there are modal facts commits one to the position known as essentialism, which is examined in the fifth chapter. Recent versions of essentialism fail to provide a ground for modalities because they define essential properties as necessary properties. This immediately succumbs to a charge of circularity. What is required is a theory that defines necessary properties in terms of essential ones. The most satisfactory account of this sort is in Aristotle.

For Aristotle, and Plato before him, the essence of a thing is found in the definition of its kind. This means that only some of the necessary properties of a thing comprise its essence. His conception of essences can provide a basis for modalities, however, because the essence places constraints on the other properties that a thing can have. Some properties, although not part of a thing's essence, will be implied by the essence, in the sense that any object that has that essence will also have those other properties. The essential properties and those that are implied by the essence will provide the basis for necessity. Possibilities will be defined in terms of properties that are neither implied by the essence nor contrary to the essence. Knowledge of modal facts requires knowledge of essences, which in turn is provided by scientific study of natural kinds. Thus, essentialism provides a solution to both the metaphysical and epistemological problems of necessity.

The position that things have essences has certainly been controversial. The final chapter examines two of the most important objections to essentialism. The first is the problem of species essentialism. A good candidate for an essential property of an organism is its species. Biologists have argued, however, that current theories of taxonomy show that species membership cannot be essential. The second. and most important objection is that an essence is relative to the description of a thing, which, in turn, is relative to our particular interests. These objections, al-though certainly worthy of consideration, can be met.

Essentialism has not enjoyed a particularly good reputation among contemporary analytic philosophers. It has, at various times, been labeled as unintelligible, perverse, and the primary reason for the intellectual stagnation of the medieval period. W.V.O. Quine, for instance, has famously called it a "metaphysical jungle." But jungles are often much more interesting than deserts; and if I am correct, then it is a jungle to which we are well advised to go.

9

CHAPTER 2

MODALITIES: FROM LOGIC TO METAPHYSICS

Introduction

We have seen that, intuitively, there is more to a sentence than simply being true or false. Its truth can be a matter of necessity, impossibility, or contingency. These features of sentences will also be features of the propositions that are expressed by the sentences. Aristotle and the Medievals considered these three features to be modes or ways that propositions are related to truth.

Most philosophers have thought that propositions do in fact have these modal features. The important questions concerning modalities, however are these: Which truths are necessary (and which are contingent or impossible)? What makes one truth necessary and another simply contingent? Here, the agreement ends.

Since this work is primarily an attempt to answer these questions. it will be helpful to first examine what others have said about them. Therefore, this chapter will consist of a short survey of the history of modal theory, from Aristotle to the development of modern modal logic. My primary aim is to show that the current primary modal theory, possible worlds theory, develops naturally from the semantics of modal logic.

The Ancient Period

Aristotle

The groundwork for all subsequent philosophical thought about modalities was laid by Aristotle. Aristotle was the first to assert that a proposition had a modal status, and that this status was something altogether different from simple truth and falsehood: "The impossible and the false have not the same significance...To say that you are standing when you are not standing is to assert a falsehood, but not an impossibility...To say, however, that you are at once standing and sitting, or that the diagonal is commensurable, is to say what is not only false but also impossible (*Cael*. I.12, 281b2-13).

Aristotle was also the first to explore the relations between necessity and possibility, and to formulate some important principles of reasoning about them:

"I use the terms 'to be possible' and 'the possible' of that which is not necessary but, being assumed, results in nothing impossible. We say indeed, homonymously¹, of the necessary that it is possible." (An. Pr. I.13, 32a18-20) "At the same time it is clear that if, when A is, B must be, then. when A is possible, B also must be possible." (Metaph. Θ 4, 1047b15-30) "Therefore it is possible that a thing may be capable of being and not be, and capable of not being and yet be, and similarly with the other kinds of predicate." (Metaph. Θ 4, 1047a21-22)

In these passages, one can find the following four principles of Aristotelian

modal logic:

L1: That which is necessary is also possible.

L2: Anything that implies an impossibility is itself impossible.

¹ In Chapter 1 of the *Categories* (1a1-3), Aristotle states that two things are homonymous if the same word, but with different senses, applies to both. Here, Aristotle uses "possible" to refer both to logical possibilities (which includes necessities) and to contingencies.

L3: Whatever follows from a possibility is itself possible.

L4: A contingent proposition may be either actually true or actually false.

L1 is intuitively true. If it is necessarily true that a triangle has three sides, then of course it must be possible for a triangle to have three sides.

L2 is the basis of the *reductio ad absurdum*, a particularly important principle in both philosophical and ordinary reasoning. We often argue that something cannot be true, for if it were, then a resulting contradiction would also be true. Since contradictions cannot possibly be true, then neither can the proposition in question. This principle provides Aristotle a method for testing possibilities. A proposition is possible only if nothing impossible follows from it.

Aristotle's argument for L3 (unlike many of his other arguments) is simple and clear: Take any possible proposition p, and another proposition q, such that qfollows from p. Assume that q is impossible, then, from the second principle above, p must also be impossible. This, however, contradicts the stipulation that p is possible. Therefore, the assumption must be false, and q is also possible.

L4 states that the contingent modal status of a proposition has no bearing on its actual truth value. Simply knowing that a proposition might be true or might be false is not enough to know whether it is actually true or false.

In chapter 12 of *De Interpretatione*, Aristotle discusses the contradictories of modal statements and the scope of the modal operators, both significant issues. Generally, if one negates a statement, the placement of the negation is not particularly crucial to the meaning of the sentence. Consider the statement, 'Bill Clinton is from Texas.' The negation of this statement is 'It is not the case that Bill Clinton is from Texas,' but the sentence 'Bill Clinton is not from Texas' is equivalent.

Sentences with modal operators, however, change meanings as the 'not' occupies different positions. Consider this sentence: 'It is possible that Bill Clinton is from Texas.' Intuitively, this is true. Clinton's parents could have moved to Texas before he was born. Is the negation of this sentence 'It is not possible that Bill Clinton is from Texas,' or is it 'It is possible that Bill Clinton is not from Texas'? These two cannot have equivalent meanings, since the first seems false; but the second is true, since Clinton is in fact not from Texas. Since contradictories cannot both be true, the negation of the original statement must be 'It is not possible that Bill Clinton is from Texas,' which is exactly what Aristotle concludes.

A particularly important distinction that Aristotle draws is the one between *absolute necessity* and *relative necessity*. One thing can be necessary relative to another thing, but not be necessary in itself. Aristotle's example is a valid argument where, "One might show by an exposition of terms that the conclusion is not necessary without qualification, though it is necessary given the premises." (An. Pr. I.10, 30b32-33) This informal notion of logical validity serves as a good example. We say that an argument is valid if and only if it must be the case that, if the premises are true, then the conclusion is true.

Consider this disjunctive syllogism. Either a democrat wins in 2000, or a republican does. A democrat does not win. Therefore, a republican wins in 2000. If this argument is valid, then the conclusion necessarily follows from the premises.

13

That does not mean, however, that the conclusion itself is necessarily true, only that it is impossible for the premises to be true and the conclusion false.

Relative necessity, then, is the absolute necessity of a conditional statement. The necessity of the conditional does not, in itself, imply the necessity of the consequent. Although Aristotle made the distinction between absolute and relative necessity, he never states that one notion can be defined in terms of the other (Kneale and Kneale 1962, 94).

The most perplexing feature of Aristotle's discussion of modalities is that he often speaks of possibility and necessity in temporal terms. In some cases, this presents no difficulty. For instance, he clearly holds the following temporal principles:

T1: If at some time it is the case that p, then it is possible that p.

Corresponding to this is

T2: If it is necessary that p, then it is always the case that p.

The first is obviously a truism. If it is true at some time that a cat is sleeping on my bed, then it must be the case that it is possible that a cat sleep on my bed. T2 follows from T1 in this way: If T1 is true, then so is it's contrapositive: if it is impossible that p, then it is never the case that p. Let q be the proposition 'p is false.' So, if it is impossible that p is false, then it is never the case that p is false. But to say that it is impossible that p is false, is just to say that it is necessary that p is true; likewise to say that it is never the case that p is false, is just to say that it is always the case that p is true (Waterlow 1992, 1). Phrasing a modal principle in temporal terms is certainly not common now, but these two principles should provide no hesitation upon reflection. What is surprising however, is that Aristotle also holds the converse of both principles:

T3: If it is possible that p, then at some time it is the case that p;

and

T4: If it is always the case that p, then it is necessary that p.

Note that each of these two principles implies the other.² It certainly requires no great reflection to note also that both are entirely implausible. It certainly seems possible that the 1996 United States Presidential election was won by a Republican. Yet, to think that therefore it is at some time true that the 1996 United States Presidential election was won by a Republican seems patently absurd.

T4 suffers a similar fate. Imagine that God has always created physical objects, and will also always continue creating physical objects. So, there is no time in which the proposition 'physical objects exist' is false. However, couldn't God have decided not to create any physical objects, or, at some time, simply destroyed every physical object? Then, it does not seem necessary that physical objects exist. Unfortunately, it is clear that Aristotle holds both T3 and T4.³

² If T3 is true, then so is 'If it is possible that $\neg p$, then at some time it is the case that $\neg p$.' The contrapositive of this sentence is 'If it is never the case that $\neg p$, then it is not possible that $\neg p$.' This implies 'If it is always the case that p, then it is necessary that p.

³ There are many instances in which Aristotle explicitly affirms one or both of these principles. They include *Top.* II.11, 115b17-18; *Ph.* IV.12, 221b28-222a9; *Ph.* III.4, 203b29-30; *Gen. Corr.* II.11, 338a1-3; *Metaph.* Θ IX. 8, 1050b7-8, and 1050b20-21; *Metaph.* N. 2, 1088b23-25; and *Cael.* I. 12, 281b25, and 281a29-30. There are also passages in which Aristotle apparently denies the principle: *Metaph.* B6, 1003a2, and A2; *Int.* 9; *An. Post.* I.6, 75a31-35; *Ph.* III.1, 200b26; and *Top.* IV.5. 156a34ff. Hintikka shows that these passages. when taken in context, do not deny the principle (Hintikka 1973, 97-102; and 1981, 60-63).

T3 is what Arthur O. Lovejoy called "The Principle of Plenitude," that every possibility is actualized at some time (Lovejoy 1936). From this, it follows that the impossible is what is never actualized. Hintikka points out that the Principle of Plenitude effectively reduces modal logic to a standard logic with tense operators (Hintikka 1981, 71-72). It is important to note that Aristotle does not define necessity and possibility in temporal terms. The four logical principles discussed above are not dependent on the temporal principles. The temporal principles, however, continued to play an important role in philosophical thinking about modality well into the modern period.

Although Aristotle successfully contributed several valuable insights into the nature of possibility and necessity, his system of modal logic is notably less successful.⁴ In *Posterior Analytics* I, he gives several examples of both valid and invalid modal syllogisms. Unfortunately, it is far from clear why certain arguments are deemed valid while other arguments, quite similar in form, are rejected. Aristotle's pupil Eudemus was the first to recognize this problem. Historians of logic have struggled with it since, and have yet to find an adequate explanation (Kneale and Kneale 1962, 89).

In *Prior Analytics* I.15, Aristotle claims that the following argument form is valid: "Let A be possible for every B, and let B belong to every C. Since C falls under B, and A is possible for every B, clearly it is possible for every C also. So a per-

⁴ There have been several analyses of Aristotle's modal syllogistic logic using contemporary logic. For example, see Lukasiewicz (1957), McCall (1963), and Patterson (1995).

fect deduction results." (33b33-35) An evaluation of this argument reveals the most

important weaknesses in Aristotle's system of modal syllogisms.

One way to schematize this argument is as follows:

(1) Possibly, all B are A.

- (2) All actual C are B
- (3) Therefore, possibly, all C are A

Charity requires admitting that this argument is very plausible at first glance. A counter-example would require a case in which B, when it becomes an A, ceases to be a C. For example, imagine the possibility in which only children can legally drink alcohol. Then, we have:

- (1) Possibly, all legal drinkers are children.
- (2) All adults are legal drinkers
- (3) Therefore, possibly, all adults are children

Converting this into a valid argument form requires a modal operator for the

second premise:

- (1) Possibly, all B are A.
- (2) Necessarily, All C are B
- (3) Therefore, possibly, all C are A

This effectively rules out any case of a B ceasing to be a C when it becomes an A, because it is now impossible for B to cease being a C.

If Aristotle were to assert a sentence of the form 'All A are B' he would also be likely to assert a corresponding sentence 'Necessarily, all A are B.' An example is the sentence 'All dogs are animals.' Nevertheless, it seems patently false that all true sentences of that form are true necessarily. Hintikka has argued that Aristotle is committed to the view that such sentences are necessary truths. Remember that Aristotle is committed to T4: 'If it is always the case that p, then it is necessary that p.' If Aristotle believes that statements of the form 'All A are B' are always true, then from T4, he must also believe that they are necessary truths. It is highly plausible that some statements of that form are always true, but surely not all of them. Aristotle, however, does believe that they are always true, for they are not to be qualified with respect to time (*An. Pr.* I, 34b7-18). So, if it is true that all A are C, then it is true of all A and of all C past, present, and future. Given this, from T4, it is necessarily true (Hintikka 1973, 136-137).

So, it appears that Aristotle, in thinking that the above argument is valid is being consistent with principles that he already holds. Unfortunately, however, this raises still another question. If Aristotle is committed to the view that all statements of the form 'All A are C' are necessary, then why does he sometimes qualify them with 'necessarily'? Either Aristotle is simply being redundant, or the necessity operator contributes something new and important to the meaning of the sentence.

There is another way to construe this argument that allows Aristotle to maintain that the second premise is non-modal. Here, one appeals to the ambiguity of the scope of the possibility operator in "Let A be possible for every B." One way to understand this statement is that there is a possible circumstance in which all B are A, and everything that is a B in that circumstance is also an A. Another way to construe this, however, is that for anything that is actually a B, there is a possible circumstance in which it is an A. The second reading has this form:

- (1) All B are possibly A.
- (2) All C are B
- (3) Therefore, all C are possibly A

This argument is obviously valid. In his discussion of the scope of the possibility operator in *Prior Analytics* I.13, 32b25-37, Aristotle clearly recognizes the distinction between 'It is possible that all A are B' and 'All A are possibly B'. Unfortunately, it is far from clear which reading of "Let A be possible for every B" he would prefer.

The examination of this syllogism reveals the two greatest problems with Aristotle's modal logic. First, without the aid of symbolization, it is extremely difficult to unambiguously specify the scope of the modal operator. Second, Aristotle's logic reflects his commitment to principles like the Principle of Plenitude that are highly implausible. Aristotle's successor, Theophrastus, improved the system somewhat, but still failed to adequately address these two problems.

Aristotle, however, deserves great credit for being the first person to seriously struggle with the issues concerning modal logic. First, he recognized that there is a significant difference between a statement's being necessarily true, it's being true in fact, and it's being possibly true but actually false. Our informal reasoning often takes these distinctions into account; therefore, our formal rules of valid inference should also.

Second, we cannot discuss the nature of logic itself without using the concepts of possibility and necessity. From Aristotle's time forward, we have defined the notion of validity in such terms. Without some clear understanding of what it means for something to be necessary, we have no clear understanding of what it means for an argument to be valid (Kneale and Kneale 1962, 96).

19

Megarians

In the later ancient period, Aristotle's logic is developed by the Peripatetic school. The Stoics provided the primary alternative during this time, using a logic that was derived from the Megarians, who, in turn, were influenced by the Eleatics, Parmenides and Zeno. Given this influence, it is not surprising to find that the Megarians are responsible for some of the more famous logical paradoxes, such as paradoxes of self-reference, referential opacity, and vagueness (Kneale and Kneale 1962, 114-115).

Diodorus Cronus was the Megarian who provided the most extensive treatment of modalities. Boethius reports that Diodorus defined the modal operators in this way: The possible is that which either is true or will be true. The impossible is that which is now false and will not be true. The necessary is that which is true and will not be false. Finally, the non-necessary is that which is either now false or will be false (Kneale and Kneale 1962, 116).

There are two things to notice about these definitions. The first is that Diodorus restricted modality to things that have a truth value, although he does not specify what he considers the proper bearers of truth to be. The second is that Diodorus, like Aristotle, conceived of necessity and possibility in temporal terms, and therefore accepted some version of the Principle of Plenitude. Every possibility is either realized now, or will be realized in the future. Unlike Aristotle, Diodorus attempted to make a positive case for the temporal nature of modality.

Diodorus' position has some odd consequences for the modal value of tensed statements. The statement 'The Titanic sunk' was possible before April 15, 1912. The same statement, however, became necessary on that date (after the time that the ship sank). Before it sank, the sentence was false, but would later be true. Afterwards, it was true, and would never be false. Likewise, the statement 'The Ti-tanic did not sink' is now impossible.

Diodorus is credited with formulating the *Master Argument* in order to prove his account of possibility. He maintained that the following three statements form an inconsistent set:

(1) Everything that is past and true is necessary.

(2) The impossible does not follow from the possible.

(3) What neither is nor will be is possible.

He then asserted that the first two are obviously true (most would grant only the incontrovertibility of the second statement), and therefore rejected the third. It is certainly not clear how he reasoned from (1) and (2) to the negation of (3). The most plausible interpretation is to assume that some sentence S is now false and always will be false. It must then be true that there was a time t at which S is false at all later times. Since this is a past truth, it follows from (1) that it is necessarily true, as is everything that follows from it, including the statement that S is now and always will be false. So, if it is necessarily true that S is false, then S is necessarily false, or impossible.⁵

Philo of Megara, a contemporary of Diodorus, defined possibility and necessity in non-temporal terms. He defined the possible as that which, when considered in itself, can be true. The necessary is that which is true, and, when considered

⁵ For an excellent discussion of the Master Argument, see Denver (1981).

in itself, cannot be false. The non-necessary is that which, when considered in itself, can be false. Finally, the impossible is that which can never be true (Kneale and Kneale 1962, 122).

The phrase 'when considered in itself' shows that Philo's notions of possibility and necessity are not relative, but absolute. Later Stoic definitions of the modal terms follow Philo, but allow for relative possibility and necessity. For example, something is possible as long as the accompanying circumstances do not prevent it in any way.

Philo's definitions seem to capture the standard contemporary conceptions of the relationships between possibility and necessity, although we have very little text upon which we can rely. One should notice here that Philo clearly takes possibility as a primitive term. To say that the possible is that which can be true is just to say that the possible is possibly true. Once one modal operator is taken as primitive, the rest can be defined in terms of it. For example, one can define the necessary as that which is not possibly false.

Diodorus and Philo presented two alternative accounts of modality. These alternatives are more than simply historically interesting. They represent the beginning of a debate that continues today. Philo's account successfully captures our intuitions about the relationship between necessity and possibility, but one could fairly object that his definition of possibility is circular. Diodorus, on the other hand, with his temporal account of modality, presents a non-circular account of possibility. The possible is not defined as what can be true, but as what is, or will be, in fact true. Diodorus' account, although non-circular, fails to capture our intuitive conceptions of necessity and possibility. If a successful non-circular account of modality cannot be found, then we must, as did Philo, simply take one modal notion as a primitive.⁶

The Medieval Period

The popular conception of medieval philosophy and logic is that it consisted simply of series of commentaries on the works of earlier philosophers, namely Plato and especially Aristotle. This is true to some extent, but it fails to acknowledge both the critical nature of those commentaries and the extent to which medieval philosophers contributed to philosophical thought.

For instance, medieval philosophers struggled with Aristotle's conception of a true statement. In the *Metaphysics*, Aristotle claimed that truth requires that the subject and predicate be combined in the thing. Medieval philosophers recognized that this account of truth is inadequate, since it fails to explain the truth of certain modal statements. On this account, necessary truths are true. For instance, the statement 'All prime numbers greater than two are odd' is necessarily true, and anything that meets the conditions in the subject (being a prime number greater than two) also meets the conditions in the predicate (being odd). Likewise, impossible statements are false.

The problem is accounting for the truth of statements about unrealized possibilities. Consider the statement 'It is possible that the capital of Oklahoma is

⁶ This distinction between Diodorus and Philo continues to be important in contemporary theories of modality. David Lewis' primary objection to Alvin Plantinga, Robert Stalnaker, R. M. Adams, et al. is that their theories are circular.

Guthrie.⁷⁷ It is false that the capital is Guthrie, but it seems true that it could be Guthrie. The subject and predicate, however, are not combined in any existing thing.

The Medievals recognized this difficulty, and the need to modify Aristotle's conception of truth to address such cases. Jean Buridan said, "This is true: Something which never will be can be', not because things are as the proposition signifies, but because things can be as the proposition signifies they can be." He explicitly affirms that there are unrealized possibilities, and even possibilities that will never be realized. Hence, he denies the Principle of Plenitude (Broadie 1993, 57-58).

The Medievals also carefully analyzed how changing the scope of the modal operator could affect the truth value of a sentence. When a whole statement was within the scope of a modal operator, the resulting statement was called a modal statement with composition. When only part of the statement was in the scope of a modal operator, it was called a modal statement with division (Broadie 1993, 57-58).

Another common conception of Medieval philosophy is that it is driven largely by theological concerns. In the case of modalities, this conception appears well-founded. One of the most interesting theological problems discussed during the medieval period is the relationship between divine foreknowledge and human

In fact, the territorial capital was Guthrie. Considering the circumstances surrounding the move of the capital to Oklahoma City; it is likely that, had things been slightly different, the capital would still be Guthrie.

free will. Boethius' treatment of this problem was the context for the first Medieval analysis of modalities.

The problem is that divine omniscience appears to be incompatible with human free will. If God is omniscient, then God knows all truths, including truths about future events. Imagine that I am now, in March, deciding where I should vacation in July. I have decided to go either to the mountains or to the beach. Assume that in July, I choose to go the mountains. If God knew in March, however, that I would choose the mountains in July, then it seems that my choice cannot be free. One condition for knowledge is truth. Since this follows from the definition of knowledge⁸, it must be a matter of necessity. So, if God knows p, then p must be true. If God knows that I will choose the mountains, then I cannot do otherwise, for then God would have been mistaken. So, if God knows what I will do in the future, then it is impossible for me to do otherwise.

Boethius' response is that this argument commits a scope fallacy. It fails to distinguish the difference between 'It is necessary that if p, then q' and 'If p, then necessarily q.' It follows from the definition of 'bachelor' that a bachelor is unmarried. So, it is necessary that if John is a bachelor, then John is unmarried. This is just to say that it is impossible that John be both a bachelor and married. It does not follow, however, that if John is a bachelor, then John is necessarily unmarried. For then, if John is in fact a bachelor, the statement 'John is unmarried' would be nec-

⁸ Any plausible definition of knowledge must restrict the class of things that are known to the class of things that are true

essarily true, and it would be impossible for John to ever be married. Surely, however, all those who are now married were once bachelors.

Likewise, since truth is a necessary condition for knowledge, it is necessary that if God knows p, then p is true. It does not follow, however, that if God knows p, then p is necessarily true. God's knowledge of p entails only the truth of p, not the necessity of p. God's knowledge of p is compatible with its being a contingent truth.

Although Boethius began to make the distinction between necessity with composition and necessity with division, he failed to completely grasp it. This can be seen in his distinction between 'simple necessity' and 'conditional necessity'. His example of simple necessity is the necessity that all humans are mortal. His example of conditional necessity is that, if you know someone is walking, then it is necessary that he is walking (Marenbon 1998a, 21). Both of these statements, however, can be expressed with the same form: 'All humans are mortal' and 'All walking people are walking.' So, if one is an example of simple necessity and the other an example of conditional necessity, then those must be different types of necessity. They cannot result simply from different placement of the necessity operator.

Anselm was the first Medieval philosopher to systematically study modality. Once again, the discussion is stimulated by a theological issue. His work, *Philosophical Fragments*, is a dialogue between a student and a teacher, presumably Anselm himself. The student raises the question of the impossibility of God's lying, and the necessity of God's acting justly. If something is impossible for God, then God is powerless in that respect, and if something is necessary for God, then God is compelled in that respect (24:16). The rest of the work (which is even less clear than Anselm's famous *Proslogion*) is an attempt to show that necessity is a causally efficacious external constraint, and possibility is a capacity to act along with the lack of external constraint (Serene 1981, 143). God, presumably, has only internal constraints, so the terms necessary and impossible do not apply with respect to God's actions.

Peter Abelard began writing shortly after the death of Anselm. Abelard was the first logician to clearly distinguish modal statements with composition and those with division, which he termed *expositio de sensu* and *expositio de re*. In the case of *expositio de sensu*, the modal operator is a predicate that has an entire non-modal sentence as its subject. The *de sensu* interpretation of "It is possible that the sitting man stand" is "'The sitting man stands' is possible." In the *de re* interpretation, the subject is not a sentence, but 'the sitting man.' So, here, the sentence is construed as attributing the possibility of standing to a particular person.

This doesn't appear to be driven by theological concerns, but it does help Abelard address some theological issues. He used this distinction to analyze Cicero's argument against the omniscience of God. Cicero had argued that contingent facts are inconsistent with the existence of an omniscient being. Since it is obvious that there are contingent truths, there can be no omniscient being. If it is possible that things were different, then it is possible that God be wrong. Presumably, it is possible for things to be different. It follows, then, that it is possible that God be wrong. It is impossible, however, for an omniscient being to be wrong. Therefore, God cannot be omniscient.
Abelard claimed that there are two ways of interpreting the crucial premise. First, on the *de sensu* interpretation, it reads, 'It is possible that, if things are different, then God is wrong.' In this case, he points out that the antecedent of the conditional is false since things are not different. He then rejects the entire conditional as false.⁹

The *de re* reading is 'If it is possible that things are different, then it is possible that God is wrong.' He grants that the antecedent is true, but denies in this case that the consequent is. If things were different, then God would know them as they were. If things were different, then God's knowledge would be also (Knuuttila 1981, 186).

Abelard's analysis of possibility certainly differs from most contemporary -analyses. Abelard thought different simultaneous possibilities were impossible. So, the *de re* interpretation of the statement 'It is possible that the sitting man stand' is either true or false depending on how one interprets it. If it is interpreted as 'The man is sitting now, and it is possible that he is standing now', then it is false. If it is interpreted as 'The man is sitting now, and it is possible that he is standing at some other time', then it is true.

One particular point that has troubled scholars is Abelard's contention that it is possible for a person whose legs have been amputated to walk. According to Marenbon, this shows that Abelard analyzed possibility in terms of what is possible for a thing, which is determined by the nature of thing, not it's specific circumstances.

⁹ Abelard's logic of conditionals is interesting. He seems to require both strict implication and relevance (Marenbon 1998b, 157).

He takes the capacity to walk to be part of the nature of the human. Therefore, it is possible for a person with no legs to walk, although that possibility will never be actualized (Marenbon 1998b, 158). If this is the case, then Abelard broke from past philosophers in the Aristotelian tradition by denying the Principle of Plenitude.

Knuuttila, however, maintains that Abelard still holds an Aristotelian, temporal view of modality, and is thus committed to the principle (Knuuttila 1981, 180). If so, then it must be the case that the person with no legs walks at some point in time. This point, presumably, is in the past.

There are, however, other claims that Abelard made that are apparently difficult to reconcile with the Aristotelian model. For example, Abelard stated that it is possible for Socrates to be a bishop. This is something that certainly never occurred at any point in time. So, if a proposition is possible if and only if it is true at some time, then it seems to be impossible that Socrates be a bishop.

To solve this problem, Abelard again appealed to nature as the source of possibility. Something is possible just in case it does not contradict nature. So, necessity is what nature demands, possibility is what nature allows, and impossibility is what nature prohibits. Abelard's epistemology is Aristotelian; one gains knowledge of nature by abstracting universal concepts from sense experience. One can only know what nature allows, however, by experiencing what nature has in fact allowed. Knowledge of possibilities requires knowledge of what has in fact occurred in things of the same nature (Knuuttila 1981, 182).

We know, then, that it is possible that Socrates be a bishop because we know that things of the same kind as Socrates have been bishops. Does this imply that Abelard is committed to the Principle of Plenitude? In order to answer this, one must first distinguish the metaphysical issue from the epistemological one. The metaphysical issue is determining the conditions under which statements of possibility and necessity are true. The epistemological issue is similar but significantly different; it is determining the conditions under which we can know that such statements are true.

So, the question then is whether Abelard intended to express conditions for truth or conditions for knowledge. Unfortunately, the text is not clear. Notice, however, that in either case. Abelard denied the traditional Principle of Plenitude. If he meant to express conditions for knowledge, then the possibility does not have to be actualized in order for it to be true. Knowing that something is possible for a thing simply requires knowing that the possibility has been actualized in something else of the same kind. I find this position to be quite plausible.

If Abelard meant to express truth conditions for possibility claims, then he still denied that every possibility must be actualized at some time. Instead, he only required that a similar possibility be actualized at some time, making Abelard a precursor of David Lewis' modal realism. So, if Abelard held the Principle of Plenitude, then he held a version that is considerably weaker than the traditional one.

Discussion of modalities in the 13th century primarily centered around the problem of God's foreknowledge and human freedom. Treatments of the problem were essentially identical to Boethius', with the important issue being that of the

distinction between modality with composition and modality with division.¹⁰ The terms *'necessitas consequentiae'* and *'necessitas consequentis'* ('necessity of the consequence' and 'necessity of the consequent', respectively) were used to clarify the distinction. It is also during this period that the term *'de dicto'* replaces Abelard's *'de sensu'*,¹¹

By this point, the doctrine that every possibility would be actualized at some point in time was being questioned. Once again, the concern was primarily theological. Surely, God, being absolutely free, could have created the world differently. God also should know those ways that the world could have been created. Since God has an idea of everything that he knows, he has ideas of possibilities. These possibilities have no independent existence. God knows possibilities simply because he knows what he can do. These are, nonetheless, possibilities that will never be realized (Boland 1996, 231).

The greatest advances in Medieval philosophy of modalities were in the 14th century with Duns Scotus and William of Ockham. The primary problem with modal theory before the 14th century was its commitment to the Principle of Plenitude, resulting in the definition of possibility as actuality at some time.¹² This, in turn, led to a temporal theory of modality: possibility is what sometimes occurs, necessity is what always occurs, and impossibility is what never occurs. Duns Scotus was the

¹⁰ For example, see Aquinas' Summa Contra Gentiles I. 67. Aquinas states that necessity de dicto does not imply necessity de re. The Kneales claim that this passage shows that he did think implication held for the converse (1962, 237). I find no evidence for that in the text. however.

¹¹ The term first appears in *De Modalibus*, a tract that is attributed to Aquinas, but is not universally accepted as genuine (see Bochenski 1940).

¹² Although, as we have seen, this definition is at least weakened by the time of Abelard.

first to clearly reject this account of possibility in favor of what is now called logical possibility.

The problem that Scotus saw with the temporal theory of modality is that it entails that everything that occurs, at any particular moment in time, occurs necessarily, because there is only a single possibility for any given time. Scotus rejected this, stipulating that possibility requires that something else could have happened at the very same time as the actual event. So, even though something is always true, it may be only a contingent truth, and it is something that might be possibly true even though it is actually never true. For example, consider the sentence "Kennedy was shot on November 22, 1963." This certainly seems to be only contingently true. Since it is true, however, there is no time at which it is false. So, on the temporal account of modality, it is necessarily true. Likewise, the sentence "Kennedy was poisoned on November 22, 1963" is necessarily false.

One of Scotus' purposes in rejecting the temporal account was to defend freedom of the will (Scotus 1994, 27). On the temporal theory, even if a person could have willed to do something other than what she actually willed, nothing else could have occurred at that time. Any plausible account of free will, he believed, requires that something other than what actually occurred truly be possible. This, in turn, enabled him to address the question of God's foreknowledge. Even though God has always known what the future holds, that knowledge is contingent. God would have known something else, had that other possibility been actualized.

This account is really an account of logical possibility, which for Scotus was anything that is not a contradiction. Not all logical possibilities are real possi-

bilities, however. The structure of the natural world places limits on the range of real possibilities. Every real possibility must at least be logically possible, so a theory of modality must at least have some account of logical possibility.

Every statement that is not a contradiction is logically possible. Each of these logically possible statements can be conceived of as being actually true. Scotus admitted that not every possibility can be actual, because not all possibilities are *compossible*, meaning that they cannot be actualized together. Possibility, then, is conceptual consistency. Scotus divided the realm of possibility into different classes of compossible states of affairs, the actual world being one of those. It is here that we find the foundation of 20th century possible worlds theory (Knuuttila 1982, 353-355).

In Summa Logicae, William of Ockham developed the ideas of Scotus into a systematic modal logic. Ockham's first task was to formulate rules for valid modal inferences. As in most early logic texts, this was done by giving examples of both valid and invalid arguments. For example, he states that the inference from 'If p, then q' to If p is necessary, then q is necessary' is invalid. He formulated rules for sentences including 'necessary', 'possible', 'contingent', and 'impossible'. He also formulated rules for arguments that contained different types of modal claims within the same argument (Ockham 1988, 314-319, 321-327).

Interestingly, he recognized the relationship between modal operators such as 'necessary' and other non truth-functional operators such as 'know' and 'believe'. Ockham maintained that the *de re* and *de dicto* forms using such operators were equivalent if the subject was a proper name or demonstrative pronoun. His

examples are the statements 'This is known to be white' and 'It is known that this is white.' If the subject is neither a proper name nor a demonstrative, however, they are not equivalent: 'A white thing is known to be a man' is not equivalent to 'It is known that a white thing is a man.' In the latter case, what is known is that the statement 'A white thing is man' is true. In the former, what is known is that some object, which happens to be white (although this may not be known), is a man (Ockham 1988, 319-320).

The next section of Ockham's work concerns the relationships that the modal operators bear to one another. First, he listed the incompatible modalities, such as 'necessary' and 'contingent'. Second he listed pairs made up of "logically superior and inferior modalities." These are one-way entailment relations, such as 'everything that is contingent is possible.' Third, are the 'independent' modalities, such as 'possible.' Fourth are the jointly exhaustive modalities, such as 'possible' and 'impossible.' Then, he listed modalities that are incompatible with more than one other modality, such as 'contingent', which is incompatible with both 'necessary' and 'impossible'. He then discussed the equivalence relations between modal sentences. One example is 'It is not possible that every animal is a man' is equivalent to 'It is necessary that some animal is not a man.' (Ockham 1988, 328-336)

At this point, it should be clear that our current intuitions involving the relationship between possibility and necessity were well in place by the end of the 14th century. Even without the aid of the predicate calculus, Ockham had successfully systematized much of contemporary modal logic. In fact, there were no real ad-

vances in modal logic for five hundred years, although important philosophical study of modalities continued.

An important issue during the Medieval period was the relationship of modal truths to the mind of God. Before Scotus and Ockham, the consensus was that necessity is somehow dependent upon the mind of God. Aquinas wrote, "Necessary means a certain mode of truth; and truth, according to the Philosopher (*Metaph.* vi), is in the mind. Therefore in this sense the true and necessary are eternal, because they are in the eternal mind, which is the divine intellect alone; hence it does not follow that anything beside God is eternal." (*Summa Theologica*, I, q. 10, a. 3, ad. 3) Necessary truths are not intrinsically eternal, but are eternal only because of their relationship to God. This includes, presumably, mathematical truths and the truths of logic: "The nature of a circle, and the fact that two and three make five. have eternity in the mind of God." (*Summa Theologica*, I, q. 16, a. 7, ad. 1)

In the later medieval period (14th through 16th centuries), this conception of necessity was rejected. The transition figure from the medieval period to the modern period was Francisco Suarez. Suarez believed this earlier model of necessity cannot truly capture what it means for a proposition to be necessary. Commenting on the earlier view, he wrote in *Disputationes Metaphysicae* in 1597:

Even this divine model should necessarily represent man as a rational animal and it could not represent man as having another kind of essence. This is based on nothing else but the fact that man cannot have another kind of essence, because a thing with a different kind of essence would not be a man. So, this necessity comes from the object itself and not from the divine model (Alanenen and Knuuttila 1988, 11).

Necessary truths cannot be dependent on the mental states of any person, because then they would simply not be necessary. So, what is necessary is not de-

pendent upon the will of God, but independent of it. Even God cannot will that some necessary truth be otherwise.

The Modern Period

The tendency, then, by the beginning of the modern period, was to separate modality from the will of God. Descartes, however, took the contrary position. In a letter to Mersenne on April 15, 1630, Descartes claimed that mathematical truths were stipulated by God, and depend completely upon God for their truth. Again, in another letter to Mersenne dated May 6, 1630, he wrote, "As for the eternal truths, I say once more that they are true or possible only because God knows them as true or possible. They are not known as true by God in any way which would imply that they are true independently of him."

On Descartes' view, if necessary truths are independent of God's will, then God's power is limited. God, being omnipotent, is subject to no restrictions, apparently not even to logical restrictions. Descartes claimed repeatedly that God can make necessary propositions false, and can also make contradictories true together (Alanenen and Knuuttila 1988, 14).

Given this, it is far from clear what conception of modality Descartes held. In fact, some doubt that Descartes held any coherent view at all (Alanenen and Knuuttila 1988, 14). The relevant passages of the text are very ambiguous. There are several possible interpretations, and each presents certain problems. One standard interpretation is that there is no absolute necessity, and everything is possible for God. Edwin Curley points out that this interpretation has serious consequences for the Cartesian view of science. The eternal truths are necessary conditions for scientific study, and as such they are immutable. Even if God created several different worlds, they would be true in all of them (Curley 1984).

One important passage occurs in a letter to Mesland (a Jesuit priest and supporter of Descartes), dated May 2, 1644: "...though God has willed that certain truths were necessary, that is not to say that he has willed them necessarily." Curley interprets this passage as a claim that the eternal truths are necessary, but they are not necessarily necessary.¹³ Alanen and Knuuttila object, pointing out that this directly contradicts what Descartes has said in other passages, that God could have made mathematical truths false (e.g. "Letter to Mersenne, May 27, 1630; and "Letter to Mesland, May 2, 1644"). So, we can come to no clear understanding of Descartes' conception of necessity (Alanenen and Knuuttila 1988, 16).

One thing that radically affected the philosophy of this period was the rise of modern science. So, although the philosophical study of modalities continued to play a role in theological thought, it also played an important role in providing a philosophical foundation for understanding science. The concept of necessity proved particularly useful in understanding the relation of cause and effect.

Thomas Hobbes primarily utilized modalities in his analysis of causation. The relation between cause and effect is a necessary relation, so that if the cause occurs, "it cannot be understood but that the effect is produced at the same instant." (Hobbes 1839, 122)

¹³ Iterated modalities quickly become difficult to understand. Alanen and Knuuttila take this to mean that God could not have made the eternal truths false, but God could have made them possibly false. If so, then Descartes' conception of modality could be formalized with neither S-4 nor S-5.

Hobbes' understanding of necessity, however, cannot support his understanding of causation. Hobbes, as with many of his predecessors, was committed to the Principle of Plenitude, the belief that every possibility would be actualized at some time. He wrote, "Every act, therefore, which is possible, shall at some time be produced; for if it shall never be produced, then those things shall never concur which are requisite for the production of it; wherefore that act is impossible, by the definition (Hobbes 1839, 129).

And so, from his commitment to the Principle of Plenitude, Hobbes concluded that every event will occur, or that every event that does occur, occurs necessarily. If the events themselves are necessary, however, there seems to be little explanatory value in postulating a necessary connection between them. If causation is necessary, and A causes B, then it is impossible for A to occur and B not occur. On Hobbes' account, however, this is trivially true, since it is impossible for B not to occur (Patoluoto 1988, 77-79).

Without doubt, the figure that has contributed the most to the contemporary philosophy of modalities is Leibniz. He is particularly known for developing the concept of a *possible world*. His other main contribution is the distinction between physical necessity and other kinds of necessity. Even so, the conclusions that Leibniz drew concerning possibility and necessity differ radically from the intuitions of most people. In order to understand Leibniz's conception of modalities, one must first understand his general metaphysical theory.

Reality, for Leibniz, consists of infinitely many individual substances, called 'monads.' God is one of these monads, and God exists necessarily, since a

contradiction follows from the assumption that God does not exist. The other monads do not exist necessarily, for no contradiction follows from their denial. So, necessity may be defined as that from which, if denied, a contradiction follows: "Indeed, a necessary proposition is one whose contrary implies a contradiction." (Leibniz 1989, 97) Likewise, a contingent truth is one whose contrary implies no contradiction.

God is omniscient, omnipotent, and omnibenevolent. The world, meaning everything that exists, was created by God (except for God himself, of course). God knew all of the various groups of compossible states of affairs that could be created; being perfectly good, God chose the best of these compossible substances to actualize. Logically speaking, however, God could have chosen a different possibility, or a different possible world (though it would have been less good and would have presumably violated the Principle of Sufficient Reason, since God would have had no reason to create a world that was less than the best). The actual world, then, which is made up of the infinitely many monads which actually exist, is simply one of an infinite number of possible worlds that God could have created (Leibniz, *The Monadology*, sec 53).

Each monad, at any given time, expresses the state of every other monad. So, if one had complete knowledge of any particular monad at some particular time, one would be able to know the state of any other monad at that time. Each state of a monad also expresses all other states, past and future, of that monad. It follows then, that complete knowledge of the state of a monad at any given time is sufficient for complete knowledge of all monads at all times; each monad is what Leibniz called, "a perpetual, living mirror of the universe." (Leibniz, *The Monadology*, sec. 56) Of course, only God is capable of such complete knowledge.

God's conception of an individual, called the complete individual concept, includes every past. present. and future state of that individual. For example, the complete individual concept of Winston Churchill includes that he was born in 1874, died in 1965, elected to Parliament in 1900, succeeded Neville Chamberlain as prime minister in 1940, and received the Nobel Prize for literature in 1953. When God created Winston Churchill, he created a substance satisfying this individual concept. Therefore, any person of whom any of this is not true is not Winston Churchill, but someone else. When considering the qualities of an individual, Leibniz did not distinguish between the accidental, those that the individual could fail to have, and the essential, those that the individual could not fail to have. All qualities have the same status. Since they could not all be accidental, they must all be essential (Mates 1986, 43).

When God considers which world to create, he is in effect simply considering which compossible group of individual concepts to actualize. There are infinitely many such groups, and each concept in those non-actualized groups reflects the others just as the actualized ones do. It is false, however, that anything at all could be actualized. Some possibilities, although possible in themselves, are not compossible with certain other possibilities, meaning that they could not both be actualized. For instance, the actual possibility that George Washington is the first President of the United States is not compossible with the possibility that someone else is. Indeed, On Leibniz's view, the possibility that someone other than George Washington were the first President is not even compossible with my existing. My complete individual concept reflects all other individual concepts, therefore, part of my individual concept is that George Washington is the first President of the United Sates. Therefore, if anyone else were, it would have been a different series of concepts that God actualized, and I would not have existed.

Each world that God could have created is maximal. It contains everything that is compossible with the other things that it contains. Also, since each individual concept reflects all of the others in the world, each concept can belong to only one world. So, the totality of individual concepts is divided into sets of individual concepts, in terms of compossibility. These sets are mutually exclusive, meaning that they have no common members. They are also jointly exhaustive, there is no individual concept that is not a member of a set. Finally, each one is maximal, there is no concept that is compossible with the members of the set but is not itself a member of the set. These sets of individual concepts are Leibniz's possible worlds (Mates 1986, 43-44).

This notion of a possible world gives Leibniz another theoretical tool with which to analyze necessity and possibility. Leibniz never stated this explicitly, but a necessary truth is one that is true in all possible worlds, a contingent truth is true in the actual world, but false in at least one of the other possible worlds (Mates 1986, 107).

The problem that Leibniz has here is easily apparent: given his notion of individual concepts and their relations to each other, how can any truth be contingent? Consider some intuitively contingent truth. For example, it seems contingent that Hank Aaron hit 755 home runs over his career. He could have hit more, he certainly could have hit fewer. If Leibniz is correct, however, it is part of Aaron's complete individual concept that he hit 755 home runs. How, then, can it possibly be false that Hank Aaron hit 755 home runs if that is part of the concept of Hank Aaron?

Another way to state the problem is that, if it is contingent that Hank Aaron hit 755 home runs, then there is a possible world in which he hits some other number. Since the possible worlds are mutually exclusive, and he exists in the actual world, Hank Aaron exists in no possible world but the actual world. So, a possible world in which Hank Aaron hits fewer or more than 755 home runs is one in which he does not exist. Therefore, such a world is one in which Hank Aaron both does and does not exist, which is a contradiction. There can be no such world. From this, it appears that Leibniz is committed to the position that every truth is a necessary truth.

Leibniz recognized this problem. In On Freedom (1689?), he stated that the reason God knows every truth infallibly is that in every true proposition, the predicate is contained in the concept of the subject. Then, he asked, "But this only seemed to increase the difficulty, for if the notion of the predicate is in the notion of the subject at a given time, then how could the subject lack the predicate without contradiction and impossibility, and without changing that notion?" (Leibniz, On Freedom, 95)

The solution that Leibniz proposed was that a contradiction could be derived from the denial of a necessary truth after a finite number of steps. A contin-

gent truth, however, required an infinite number of steps. I'm afraid that I'm in no position to explain why he thought this was a solution.

There are places, however, where Leibniz hinted at another solution. It is a contingent truth that Hank Aaron hit 755 home runs, but it is part of his complete individual concept that he do so. There are other complete individual concepts that God chose not to actualize, however, that are quite similar to the individual concept of Hank Aaron. In fact, some of them differ only by the number of home runs hit over their baseball careers (plus any other difference that is necessitated by this one, i.e. having scored more runs, etc.). So, in saying that it is a contingent truth that Hank Aaron hit 755 home runs, we are simply saying that there is a possible world in which someone very much like Hank Aaron did not hit 755 home runs.

There is at least one passage where Leibniz speaks of these alternative individuals, which later come to be known as "counterparts" (associated primarily with David Lewis). In the *Theodicy*, Leibniz wrote,

I will show you where one can find, not exactly the same Sextus whom you have seem) this is not possible, he carries always within himself what he will be). But future Sextuses who will have all that you already know of the real Sextus, but not all which, without our perceiving it, is already within him, nor consequently all that will happen to the real Sextus.

There is no evidence that Leibniz saw that this was a solution to his problem of contingency, but the strategy is certainly available (Ishiguro 1972, 123-125).

Leibniz distinguished between what he called *absolute necessity* from *hypothetical necessity*. He used this distinction to clarify the difference between logical necessity and physical necessity. Many things that were logically possible became, in effect, impossible once God decided which world to create. Given that God has created this world with these physical laws, then certain events become impossible. For instance, he thought that it is impossible in this sense that a moving object change its course unless it is acted on by some other object. Such things are hypothetically impossible, but absolutely possible. Hypothetical necessity, then, requires two conditions. The first is that a conditional having the form 'If God creates world w, then event e will occur' is necessary *de dicto*. The second is that the conditional not be necessary *de re*. In other words, the conditional must not have a necessary consequent (Mates 1986, 117-121). This, in effect, reduces physical necessity to a special type of logical necessity.

This notion of physical necessity, however intuitive, was soon questioned by those in the empiricist tradition, most notably by Hume. Hume's conception of necessity is similar to Leibniz's, a necessary truth is one from whose denial a contradiction follows. Applying that condition to several things that were believed to be necessary, like 'whatever begins to exist must have a cause for its existence' he discovers that these claims do not pass the test for necessity, and are therefore contingent truths (Hume 1978, 78-82). Likewise, no contradiction follows from the denial of the conditionals that Leibniz has deemed necessary. Therefore, it seems that physical necessity is impossible to demonstrate.

According to Hume, then, these relations between events cannot be relations of logical necessity. He concluded that events are related by psychological necessity, or necessity that is contributed by the mind: "Necessity, then, is the effect of this observation, and is nothing but an internal impression of the mind, or a determination to carry our thoughts from one object to another." (Hume 1978, 78-82, 165) This was echoed by Kant in the *Critique of Pure Reason*: "The categories of modality have the peculiarity that, in determining an object, they do not in the least enlarge the concept to which they are attached as predicates. They only express the relation of the concept to the faculty of knowledge." (Kant 1929, B266)

In the *Treatise*, Hume hinted at an even more radical conclusion, that physical necessity is really no different in this way from logical necessity. So. all necessity is psychological necessity. He wrote, "Thus as the necessity, which makes two times two equal to four, or three angles of a triangle equal to two right ones, lies only in the act of the understanding, by which we consider and compare these ideas; in like manner the necessity of power, which unites causes and effect, lies in the determination of the mind to pass from the one to the other." (Hume 1978, 78-82, 166) This is, however, the only place where Hume makes such a claim (Flew 1986, 71-72).

At the end of the 18th century, there were two competing perspectives concerning modalities. One perspective, represented by Leibniz, is that possibility and necessity play crucial roles in addressing certain philosophical issues, such as the existence of God, human freedom, and natural laws. The other perspective, represented by Hume, is that most, if not all, necessity has nothing to do with reality *per se*, but only in our perception of it, or in relations of ideas. As we shall see, this important debate continues today.

The Formal Systems of C.I. Lewis

Although the metaphysics of modality played an important role in the philosophy of Leibniz, no significant work was done in modal logic until fairly recently. The first hint of what is to become contemporary modal logic is found in the writings of Hugh MacColl at the end of the 19th century in his work *Symbolic Logic* and its Applications (MacColl 1906).

MacColl's logic included the operations of disjunction (a + b), negation (a')and implication (a : b). He asserted as a valid principle

(1) (a : b) : (a' + b)

but denied

(2) (a:b) = (a'+b)

or, in more recent notation

$$(1^*) (a \rightarrow b) \rightarrow (\sim a \lor b)$$

and

 (2^*) $(a \rightarrow b) \leftrightarrow (\neg a \lor b)$

For MacColl, then, implication and material implication are not equivalent. The former entails the latter, but the latter does not entail the former. Therefore, material implication is a weaker notion than implication. Implication must be a modal notion, which can be seen in the examples that he gave. His argument for denying (2) was that if a is the proposition 'He will persist in his extravagancy' and b is the proposition 'He will be ruined', then the negation of 'a : b' is 'He may persist in his extravagancy without necessarily being ruined', while the negation of 'a' + b' is 'He will persist in his extravagancy and he will not be ruined. MacColl denied that these two are equivalent because the first asserts only a possibility while the second asserts something stronger. MacColl explicitly denied that his implication operator is truth-functional by defining (a : b) as $(a' + b)^{\epsilon}$, where ϵ represents necessity (Hughes and Cresswell 1996, 193-194).

MacColl provided no axioms for his system, so it is impossible to determine which of the current modal systems most closely matches his. Modern modal logic really began with the work of C.I. Lewis. The development of modern modal logic awaited an axiomatic treatment of non-modal logic, which was provided by Russell and Whitehead with the 1910 publication of *Principia Mathematica* (Whitehead and Russell 1910).

In *Principia*, Russell and Whitehead derived the following theorems from their axioms:

 $(3) p \supset (q \supset p)$ $(4) \sim p \supset (p \supset q)$

If we view ' \supset ' as standing for some sort of implication, then (3) can be understood as meaning that if a proposition is true, then it is implied by any other proposition. (4) is that if a proposition is false, then it implies any other proposition. These two theorems are often called the paradoxes of material implication. Furthermore, since the Law of Excluded Middle ($p \lor \neg p$) tells us that either p or its negation must be true, then it is simple to derive from (3) and (4) this theorem:

$$(5) (p \supset q) \lor (q \supset p),$$

that is, for any pair of propositions, either the first implies the second, or the second implies the first.

Lewis did not wish to reject these theorems entirely, but only to point out that they are a result of using a material implication operator to translate the word 'imply'. Lewis further pointed out that this is not what we usually mean by 'imply'. From (3), we see that the proposition 'Julius Caesar ate a banana for breakfast on the morning of his death' implies the proposition 'Bill Clinton won the presidential election in 1996'. From (4), we find that the proposition 'Abraham Lincoln died in 1864' implies that the United States economic system will collapse in 2015. From (5), we find that of any two propositions, no matter how they are related or what they are about, one of them will imply the other.

The material implication of *Principia*, then, certainly does not capture the normal sense of 'imply'. In ordinary usage, when we claim that one proposition implies another, we are asserting some connection between them other than one that is simply truth-functional. Lewis held, as seems right, that there is some stronger sense of 'imply' that we generally mean. In this sense of 'imply' it is false that every true proposition is implied by any other proposition, nor is it the case that any proposition is implied by any false proposition. Also, in this stronger sense, there are pairs of propositions such that neither one implies the other. So, on Lewis' understanding of 'imply', (3), (4), and (5) are all false.

The sense of implication that Lewis had in mind is one he called *strict implication*, which he symbolized with the fishhook, ' \dashv '. He interpreted $p \dashv q$ to mean that it is impossible for p to be true and q false. So, the proposition 'Abraham Lincoln died in 1864' will not strictly imply that the United States economic system will collapse in 2015, since it certainly seems possible that even if Abraham Lincoln were to have died in 1864, the economic system might collapse earlier or later than 2015, or fail to collapse at all.

Lewis sometimes took strict implication to be primitive, sometimes, strict disjunction, and sometimes logical impossibility. Most contemporary modal logic systems take either logical possibility (\Diamond) or logical necessity (\Box) as primitive. All of these operators are interdefinable, so that any operator can be used without changing the system. The necessity and possibility operators are interdefinable in the same way that the universal and existential quantifiers are in first order logic. To say that *p* is necessary ($\Box p$) is just to say that it is not possible for the negation of *p* to be true. ($\neg \Diamond \neg p$).

Lewis' strict implication is equivalent to the logical necessity of the material conditional. If $p \prec q$, it is impossible for p to be true and q false. So, $p \prec q$ is equivalent to $\sim \Diamond(p\&\sim q)$, which is equivalent to $\Box(p \supset q)$. Lewis' strict disjunction can be symbolized as $\Box(p \lor q)$, and logical impossibility is either $\sim \Diamond p$ or $\Box \sim p$.

When a modal operator is added to the standard propositional logic calculus, modal propositional logic systems can be formed. An axiomatic presentation of a logical system requires three things. First, it requires a specification of the language of the system. This is a list of the primitive symbols, and a set of formation rules that show which strings of the primitive symbols count as well-formed formulas. Second, the system will contain a set of well-formed formulas, which are known as the axioms of the system. Third, it requires a set of transformation rules that allow one to derive new well-formed formulas from the axioms. The formulas obtained from using the transformation rules in this way are called the *theorems* of the system. Notice that many different sets of axioms and transformation rules are possible. and many of these are likely to result in different systems of propositional modal logic. In fact, much of the early work in modal logic was in developing different, non-equivalent systems. In his book, *A Survey of Symbolic Logic*, Lewis developed an axiomatic modal system using logical impossibility as the primitive modal operator, and conjunction and negation as truth-functional operators. Later, in the book *Symbolic Logic*, which Lewis co-wrote with C. H. Langford, there is a comprehensive treatment of that system and four others (Lewis and Langford 1959, 492-514). Altogether, Lewis developed eight different systems, and still more systems were developed by other logicians. These systems all differ according to their axiom sets.

One of the most basic systems (not developed by Lewis) is called K. It takes as its axioms all of the truth-functional tautologies and the following modal formula:

$$\mathsf{K}: \square(p \supset q) \supset (\squarep \supset \squareq)$$

K has two transformation rules: modus ponens and the rule of necessitation. The rule of necessitation says that if p is either an axiom or a theorem of K, then $\Box p$ is also a theorem, i.e. where ' \vdash ' means 'is derivable from,' then if $\vdash p$ (meaning p is derivable from just assumptions, all of which can be discharged in the derivation), then $\vdash \Box p$. Lewis rejected K, because the formula T: $\Box p \supset p$ is not a theorem of K.

Intuitively, this formula says that if a proposition is necessarily true, then it is true.¹⁴ If one adds this formula as an axiom to K, we get the system T.

The Brouwersche system contains T and adds the following as an axiom: B: $p \supset \Box \Diamond p$

S-4 differs from the Brouwersche system by adding to T, instead of B. this axiom:

S4: $\Box p \supset \Box \Box p$

The system S-5 is formed by adding both B and S4 to T. This is equivalent to adding to T the following formula:

S5: $\Diamond p \supset \Box \Diamond p$

There are other modal systems, but just these five already point to a problem. We require a logic system to tell us which inferences are valid. In a modal context, however, we have already seen five different answers to the question of which modal inferences are valid. Which system is correct?

The most obvious strategy is examine the axioms of each system and choose the one that best fits our intuitions about necessity and possibility. K seems correct. K simply says that anything that follows from a necessary truth is itself a necessary truth. T says that if a proposition is necessarily true, then it is true. Both of these seem straightforwardly true.

Intuition is not so clear a guide in the case of the Brouwersche system, S-4, and S-5, however. B claims that if a proposition is true, then its possibility is neces-

¹⁴ In most senses of necessary. such as logical necessity and physical necessity, this formula seems right. There are certain contexts in which modal logic might be used, however, in which this formula is false. An example is deontic logic, where the operators are understood as meaning moral obligation and permissibility. Just because an action is morally obligatory does not guarantee that it will be performed.

sary. S4 claims that necessary truths are necessarily necessary. S5 says that the modal status of a proposition, no matter what it is, is a matter of necessity.

Here, it is not straightforwardly clear just which of these claims, if any, are true, nor is it even completely obvious what they mean. We could judge the systems by examining the formulas that can be derived in each one, but these are even less clear than the axioms themselves. So, by developing several non-equivalent systems, none of which is obviously the right one, Lewis' work in modal logic called into question the legitimacy of modal inference itself.

The problem was that there was no systematic way to interpret what the modal formulas meant. Logicians could only appeal to the ordinary language interpretations of 'necessarily' and 'possibly.' By listing the axioms and derivation rules for each system, Lewis and his successors succeeded in specifying the syntax of various modal logic systems, but they failed to specify the semantics. Without a semantics, there can be no systematic understanding of what a true, or valid, formula is. Without this notion of a valid formula, the system cannot be proved to be complete, that is, one cannot show that the system can produce all of the modal truths.

Quine's Modal Skepticism

One might take the lack of a formal semantics to be a good reason to believe that there is nothing to modal logic besides the manipulation of a formal language. This century has certainly seen a revival of Humean skepticism concerning modalities, led notably by W.V.O. Quine, but reinforced by the empiricist spirit of the times.

Quine's attack on the concept of necessity follows from his denial of the analytic/synthetic distinction. Analyticity has been defined in various ways since the time of Kant. Kant defined an analytic truth as one whose predicate is included in the concept the subject, or as one whose negation entailed a contradiction. Later, Frege urged that a sentence is analytic if it meets at least one of two conditions: 1) it is a truth of logic, or 2) it is reducible to a truth of logic by substituting synonymous terms (Frege 1980, 4). Analyticity has recently been viewed as truth purely in virtue of meaning. Analyticity, when viewed in this way, could easily be understood as providing the grounds for necessary truth. Since an analytic truth is true solely in virtue of its meaning, it could not possibly be false. So, analytic truths are necessary truths.

"Two Dogmas of Empiricism" of 1951 contains Quine's famous attack on the notion of analyticity, specifically on Frege's account of an analytic truth as one that is reducible to a logical truth. On Frege's account, the sentence "All bachelors are unmarried" is analytic since it can be changed to the logical truth "All unmarried males are unmarried" simply by substituting the term 'bachelors' for the synonymous term 'unmarried males.'

Quine's objection focuses on the notion of synonymy that this requires. Any account of synonymy will either be an insufficient ground for analyticity, or it will beg the question. One might argue that 'bachelor' is synonymous with 'unmarried male' because the latter is simply the definition of the former. Lexical definitions,

however, are simply based on the lexicographer's understanding of synonymous usage; so such attempts to explain synonymy are inherently question-begging.

Another tactic is to appeal to substitutability *salva veritate* (with no change in truth conditions). Then, 'bachelor' and 'unmarried male' are synonymous just in case one may be substituted for another in any context with no change in the truth value of the sentence. Of course, *salva veritate* substitution in extensional contexts will never be sufficient to ground synonymy. At most, it would show that there happen to be no bachelors who are married. Substitution will only be sufficient to account for synonymy if it holds in all contexts, both extensional and intensional. So, 'bachelor' and 'unmarried male' are synonymous only if *salva veritate* substitution holds in for "Necessarily, all bachelors are unmarried" and other contexts. According, to Quine, however, a sentence is necessary if and only if it is analytic. He concludes that we can understand necessity only if we already understand analyticity (Quine 1953, 25-32). His attack on analyticity becomes a simultaneous attack on necessity.

In truth, however, this comprises only a partial attack on analyticity, and therefore only a partial attack on necessity. Even if one concedes that Quine's argument is successful; he has not shown that there are no analytic truths, nor has he attempted to do so. At best, he has shown that there are conceptual problems with the second of Frege's two conditions. Sentences that are analytic in virtue of their logical truth are still not questioned. So, the concept of logical necessity still stands. Also, a strong conclusion concerning necessity only follows given the assumption that all necessary truths are analytic. If there is a type of necessity that is fundamentally non-linguistic, Quine's argument will not count against it.

Quine's vehement distrust of necessity follows in part from his view that "No statement is immune to revision." (Quine 1953, 43) The failure of the distinction between the analytic and synthetic entails that any statement could be revised in light of future experience. So, there are no sentences that are true no matter what, and what is a necessary truth but one that is true no matter what? Upon closer examination, this principle, at best, presupposes a notion of necessity; and, at worst, is untenable.

The claim presupposes necessity in that it presupposes a notion of possibility, and, as we have seen, the two are interdefinable. By the claim that there is no statement that is immune to revision, Quine cannot mean that there is no statement that will in fact sometime be revised. Instead, this must mean that there are no statements that are, in principle, unrevisable, that is, impossible to revise in the light of future experience.

Even worse, Crispin Wright has shown that this principle is untenable, leading to a vicious regress. Quine's position is that a statement should be revised in face of recalcitrant experience. If all statements are candidates for revision, then one can revise some of the empirical claims, claims about the underlying logic, or claims about the logical consequences involved. The third option is available since we could solve the problem of recalcitrance by rejecting the claim that contradictions are bad. In fact, there may be times when we should do this, if it maximizes empirical consistency and theoretical simplicity. To decide if this is the case, however, one must make judgments concerning the consequences of the options, which requires hypotheses concerning logical consequence. In making these judgments, one should use the best hypotheses, and one determines which are the best hypotheses by applying the pragmatic criteria of empirical consistency and theoretical simplicity. So, the process is either circular or viciously regressive (Wright 1986). So, all things considered, there is now little reason to concede the success of Quine's argument against necessity.¹⁵

Kripke and Modal Semantics

In the late fifties and early sixties, several logicians developed a scheme for a semantics for the various modal logic systems. At the foundation of this scheme lay the Leibnizian notion of the possible world. Necessary truth is truth in all possible worlds. Using this notion, it became possible to understand formally and semantically exactly how the various modal systems differed from each other, and how the modal operators could be understood uniformly across different modal systems. One approach that is now often used was developed by Saul Kripke in his "Semantical Considerations on Modal Logic." (1963)

Propositional Logic

A formal language is an uninterpreted system that contains sentences constructed by recursively applying rules of construction to some basic elements of the

¹⁵ Quine has other arguments against specific uses of modalities. These arguments are primarily directed to the essentialist nature of modal logic. a topic that is discussed in later chapters.

language. In non-modal propositional logic, these basic elements are atomic sentences and logical operators. An atomic sentence is one that contains no other sentence as an element. Atomic sentences are represented by upper-case letters, A, B, C.... Complex sentences are constructed from these atomic sentences using the standard logical connectives.¹⁶ The rules that govern construction of complex sentences are:

(1) If A is a sentence, then $\neg A$ is a sentence.

(2) If A and B are sentences, then A & B is a sentence.

Once it has been determined which formulas count as sentences of the language, one can devise semantic rules that define truth in the language. First, each atomic sentence is assigned a truth value, either true or false, but not both. Second, the truth values of the complex sentences are determined by the truth values of the atomic sentences according to the following rules:

(3) $\sim A$ is true iff A is false.

(4) A & B is true iff both A and B are true.

The semantics of propositional logic enable one to meaningfully speak of the truth values of the sentences of the language. Unfortunately, it does not allow one to speak meaningfully of the possible truth values of the sentences. This requires a semantics for propositional modal logic, which Kripke, and others. provided.

¹⁶ For simplicity, I will use only the negation and conjunction operators. Propositional logic with only these two operators is expressively complete, that is, for any sentence that is expressed using the disjunction, conditional, or biconditional operators, there is a logically equivalent sentence using only the negation and conjunction operators.

Kripke defined a model structure as an ordered triple $\langle G, K, R \rangle$. K is a set of objects, G is one of the objects in K, and R is a binary relation defined over the members of K. Intuitively, K is the set of all possible worlds, G is the actual world, and R is the accessibility relation (a world w is accessible to another world w* just in case every proposition that is true in w is possible in w*). Since R is defined over K in advance, it is antecedently fixed which worlds are accessible to each other. Different constraints on R give rise to alternative modal systems. Each system has its own model structure, that is, there is a T-model structure (Kripke uses the alternate name 'M'), an S4-model structure, and so on. The various model structures differ only in terms of the relation R. There is one restriction on R for the T-model structure. R has to be reflexive; every world is accessible to itself.

A T-model is a function from the atomic sentences of T and the possible worlds to truth values. A model assigns a truth value to each atomic sentence of T in each possible world. Given the model, we can determine the truth value of each complex sentence in any possible world on that model. We need only use the following familiar recursive rules.

(1) $\neg A$ is true in W iff A is false in W.

(2) A&B is true in W iff both A is true in W and B is true in W.

(3) $\Diamond A$ is true in W iff there is at least one possible world W, accessible to W, s.t. A is true in W (i.e. iff there is some W such that WRW and A is true in W),

(4) $\exists A$ is true in W iff for every world W that is accessible to W, A is true in W.

A formula is T-valid just in case it is true in all models with a T-model structure. This means that a formula is valid just in case, when we restrict the accessibility relation only by specifying that it must be reflexive, the formula comes out true in all possible worlds under all assignments of truth value to its atomic constituents in those worlds (Kripke 1971, 63-65).

What is significant in the definition of validity in **T** is the idea that R must be reflexive, which is apparent from the axiom, $\Box p \supset p$. If $\Box p \supset p$ is false in the actual world, then the antecedent must be true in the actual world and the consequent false. That is, $\Box p$ is true in the actual world, but p is false. In order for $\Box p$ to be true, p must be true in every world accessible to the actual world. Therefore, since p is false, the actual world must not be accessible to itself. The reflexivity requirement arises from the axiom itself.

Models for the other systems impose additional structural constraints on R. R is reflexive in **B**, S-4, and S-5, since they all contain **T**. In the Brouwersche system, R is also symmetrical. The Brouwersche system adds $p \supset \Box \Diamond p$ as an axiom to **T**. In order for this to be false, p must be true, and $\Box \Diamond p$ false. For the latter, there must be a world accessible to the actual world for which there is no accessible world in which p is true. In order to prevent this, we only need to make the actual world accessible to every world that is accessible to it. This amounts to imposing a requirement of symmetry on R.

The characteristic axiom for S-4 is $\Box p \supset \Box \Box p$. In order for the antecedent to be true, p must be true in every world accessible to the actual world. In order for the consequent to be false, p must be false in a world that is accessible to a world accessible to the actual world. This is prevented by requiring transitivity for R. Since S-5 contains T, B, and S-4. R must be reflexive, symmetrical, and transitive.

First-Order Logic

Corresponding to each propositional modal system is a quantified modal system. The semantic rules for first-order languages are understandably more complex than those for propositional logic, since the basic elements of first-order languages are not atomic sentences but terms and predicates.

A first-order model for a non-modal language consists of a *domain* D and a *valuation function* v. D is a nonempty set of objects; v is a function from n-place predicates to sets of n-tuples of D. An *interpretation* I of the language is a function assigning values to all terms, predicates and sentences relative to a model.

For quantified modal logic, the model must contain not simply a domain of individuals, but a set of possible worlds, each world having its own domain of individuals. The valuation function must assign values to the predicates relative to each possible world.

To accomplish this, Kripke introduced a "quantificational model structure." It includes the ordered triple $\langle G, K, R \rangle$ along with a function ψ that assigns to each $H \varepsilon K$ (if K is a set of possible worlds, then H is some particular world) a set $\psi(H)$, which is the domain of H. This is the set of objects existing in a particular possible world (Kripke 1971, 68-67).

A quantificational model includes the model structure, which determines the domains of the worlds, and a valuation function, which is now a function from n-place predicates and worlds to sets of n-tuples from union of the domains of the possible worlds (all possible objects). These are the extensions of the predicates in the various worlds (Stalnaker 1977, 333-335).

In modal logic, the modal operators appear to function very much like quantifiers. $\Box P$ is equivalent to $\sim \Diamond \sim P$, just as $\forall xFx$ is equivalent to $\neg \exists x \sim Fx$. In order for $\forall xFx$ to be false, it only requires one thing that does not have the property F. In order for $\Box p$ to be false, it only requires one possible instance in which the proposition p is false.

Understanding the modal operators as quantifiers, however, introduces a new problem. Quantifiers in first-order logic quantify over objects. If the modal operators are quantifiers, what sorts of things do they quantify over? Given Kripke's intuitive understanding of the semantics, the natural answer is that they quantify over possible worlds.

The natural English reading that philosophers give to $\Box p$ and $\Diamond p$ lends support to this conclusion. $\forall xFx$ is naturally read as 'For all objects x, x if F.' $\exists xFx$ is read as 'For some object x, x is F.' $\Box p$ is usually read as 'p is true in all possible worlds', or 'For all possible worlds, p is true.' $\Diamond p$ would then be 'For some possible world, p is true.'

Kripke's intuitive understanding of the semantics also appeals to possible objects existing in possible worlds. This is supported by the common intuition that there could have been more things than there actually are. This notion serves as a very powerful heuristic to aid one's understanding of the various modal systems.

The possible worlds account helps us decide on a particular system. Do we think that what is possible can vary across possible worlds? If so, we commit ourselves to either T, B, or S-4; if not, then S-5. Since possibility is invariant in S-5, only it captures the Leibnizian view of necessity as truth in all possible worlds (Loux 1979, 24-28).

The problem with this apparatus is that it seems committed to possible worlds. It seems that, to make sense of modal logic, we have to suppose that in some sense there really are possible worlds other than our own and that in some sense there really are possible objects not found in the actual world. One can respond that while we informally talk of possible worlds and objects, we're only committed to the set theoretical machinery that is employed in the semantics. That is what is essential to the project.

One is saying here that the systems should be thought of as completely uninterpreted systems of inscriptions incorporating a primitive vocabulary and two kinds of rules: Rules that allow constructions of strings and rules that allow the derivation of other strings. Viewed this way, the systems are not about anything. No set of objects is privileged. Any set of objects that satisfy the formal constraints imposed by the model account can play the role of K and U. We can think of K as a set of possible worlds, but nothing commits us to that.

There seems to be something right here. Defining validity does not require these worlds. But can we really think of the modal systems this way? Some have claimed that we cannot if we are to take them as systems of modal logic that provide theories of legitimate modal inference. Then we must think of them as constituting the subject matter of our modal discourse (Menzel 1990, 361-363). In his book *Counterfactuals*, David Lewis introduced a metaphysics of possible worlds. The opening paragraph of that chapter is among the most famous in contemporary metaphysics:

I believe that there are possible worlds other than the one we happen to inhabit. If an argument is wanted, it is this. It is uncontroversially true that things might have been otherwise than they are. I believe, and so do you, that things could have been different in countless ways. But what does this mean? Ordinary language permits the paraphrase: there are many ways things could have been besides the way they actually are. On the face of it, this sentence is an existential quantification. It says that there exist many entities of a certain description, to wit 'ways things could have been.' I believe that things could have been different in countless ways; I believe permissible paraphrases of what I believe; taking the paraphrase at its face value, I therefore believe in the existence of entities that might be called 'ways things could have been.' I prefer to call them 'possible worlds.' (Lewis 1973, 84)

For the uninitiated, talk of possible worlds is mysterious. Philosophers, however, having been sufficiently initiated into the rites of modal metaphysics with all of its arcane language, use the term with apparent ease. In the passage quoted above, Lewis tries to convince the reader that he means nothing mysterious by 'possible world.' He only means a way that things could have been, and we all believe that there are many different ways that things could have been.

So, Lewis argues, we all believe in something, or rather some thing, a way that things could have been.¹⁷ This thing is a possible world. Now by claiming that these worlds are things, or "entities", Lewis recedes back into the shrouds of mystery. We seem to be in a dilemma. Apparently, we need to view possible worlds as real in order to make sense of modal claims, but how are we to make sense of the

¹⁷ Not everyone has found this argument convincing. For example see Susan Haack's excellent discussion (1977).
possible worlds themselves? What are these things like? What sorts of things are they? The next chapter examines the two basic answers that have been proposed to this question.

CHAPTER 3

POSSIBLE WORLDS: COMPETING PERSPECTIVES

Introduction

With the development of the formal semantics for modal logic, the phrase "possible world" is now very entrenched in the philosophical lexicon. One should be careful, however, not to take the wide acceptance of the term as any indication of agreement about what it means.

There are three primary types of positions concerning possible worlds. The first is a metaphysically neutral position in which the language of possible worlds simply serves as a heuristic that enables us to analyze the logical relationships between modal sentences. The phrase 'possible world' simply allows one to restate modal sentences in a more perspicuous manner. Thus, 'In all possible worlds, 2+2=4' just means 'It is necessary that 2+2=4.' Here, the phrase 'possible world' does not function logically like a term. So, the use of the phrase does not commit one to the existence of anything. One cannot, therefore use the semantics of modal logic to provide any insight into the metaphysics of modality.

This is certainly a safe position, but is also less than completely satisfying. The interesting philosophical issue concerns not the paraphrasing of modal claims, but rather their analysis. If the concept of a possible world could enable one to state the truth conditions of modal statements, then it is likely that this would have several additional theoretical benefits. Following a discussion of those benefits. I will turn to the substantive accounts of possible worlds that have been offered in recent years.

The Case for Possible Worlds

As noted in the previous chapter, a natural inclination is to interpret the semantics of modal logic as quantification over possible worlds. This inclination, combined with Quine's famous dictum "To be is to be the value of a bound variable," appears to commit one to the existence of certain objects called possible worlds. As a general rule, however, one should make as few ontological commitments as necessary, especially to things that seem *prima facie* implausible. Therefore, if the concept of a possible world were to provide no theoretical benefits, one would have good reason to believe that the deflationary account is best. The proponents of possible worlds have been quick to offer a cornucopia of benefits, so many in fact, that David Lewis has called an ontology of possible worlds "a philosopher's paradise." (Lewis 1986, ch. 1)

The first benefit, obviously, is in the analysis of modalities. Possible worlds provide a simple way to distinguish necessary, contingent, and impossible existence. Necessary objects exist in all possible worlds. Contingent things exist in some worlds but not in others. Impossible objects exist in none.

A possible worlds account also has no difficulty explaining the difference between *de re* and *de dicto* modalities. The *de dicto* claim 'It is necessary that all dogs are mammals' ascribes necessity to the sentence 'All dogs are mammals.' So, the *de dicto* claim is true just in case that non-modal sentence is true in all possible worlds. The *de re* claim 'All dogs are necessarily mammals' does not assert the necessity of a sentence, but instead seems to ascribe a modal property to a group of things that actually exist. In this case, it claims that every existing dog is necessarily a mammal, leaving open the possibility that non-existing dogs might not have been mammals. So, this modal claim is true just in case every thing that is a dog in the actual world is also a dog in every possible world in which it exists. The restriction "in which it exists" is needed, for otherwise, each actual dog is not only necessarily a mammal, but also necessarily exists.

One difficulty facing any analysis of modality is the many different uses of the words 'necessary,' 'possible,' and their variants. A cursory examination of the way these words are used in English reveals many different types of necessity and possibility. For instance, consider the following sentences:

L: It is necessary that either a sentence or its negation is true.

M: Socrates is necessarily human.

P: An object falling in a vacuum must fall at a rate of 9.8m/s^2 .

E: For all I know, it is not necessary that George W. Bush win the presidential election in 2000.

Each of these sentences refers to some different type of necessity. The first is an expression of *logical necessity*, claiming that the logical Law of Excluded Middle is a necessary truth. Most would deny that the sentence 'Socrates is human' expresses a logical truth, but many would affirm that it still necessarily true. Its necessity, then, is not logical necessity, but *metaphysical necessity*. That objects fall at a particular rate in a vacuum is not a logical truth, but instead a truth of science. So, the third sentence is not an expression of either logical or metaphysical necessity. It is instead one of *physical*, or *nomological necessity*. The fourth sentence is a truth of neither logic nor science. It is a claim about what is necessary given what I know to be the case. This is sometimes called *epistemic necessity*. Of course, these are not the only uses of 'necessary' in English. There is *doxastic necessity*, or what is necessary given what I believe to be true; *historical necessity*, what is necessary given the history of the world to a certain point in time; and *analytic necessity*, what is logically necessary if one fixes the meanings of the non-logical terms.¹ Since necessity and possibility are interdefinable, there is a type of possibility that corresponds to each of these types of necessity.

A major strength of a possible worlds analysis of modality is that it can treat these different types of necessity and possibility with equal ease. Logical possibility and necessity are simply cases of quantification over the entire set of possible worlds. The sentence 'It is necessary that either a sentence or its negation is true' is true just in case for any possible world w and for any sentence S, either S is true in w or S is false in w. A sentence S is logically possible just in case for some world w, S is true in w.

Other types of necessity and possibility are treated in the same way, except quantification is limited to a subset of the set of all possible worlds. These worlds will all be like the actual world in some particular way. It will be nomologically necessary that x falls at a rate of 9.8m/s^2 if and only if for all possible worlds having the same physical laws as the actual world, x falls at the rate of 9.8m/s^2 . An analysis

¹ Many philosophers. following Quine, will find analytic necessity problematic. Notice that the problem is not with the analysis of analytic necessity, but with the prospect of ever fixing meanings.

of sentence E will restrict quantification to those worlds in which I know nothing different that I actually do. For doxastic necessity, the domain is restricted to those worlds in which I believe nothing different. Historical necessity restricts the domain to those worlds that are exactly like the actual world up to the specified time.

Once the possible worlds theorist has this framework in place, she can apply it to various philosophical problems. An obvious place to begin is the analysis of counterfactual claims. In truth-functional logic, counterfactual conditionals are notoriously problematic. Since the material conditional in truth-functional logic is true when either the antecedent is false or the consequent is true, the contrary-to-fact antecedent of the counterfactual guarantees the truth of the material conditional. So, the conditional 'If I were wealthy, then I would be happy' is true, but so is the conditional 'If I were in poverty, then I would be happy.'

On a possible worlds analysis, however, these sentences may very well have different truth values, just as they do in English. First, the possible worlds schema allows one to discuss situations in which the antecedent is true. So, one could ask if I am happy in a possible world where I am in poverty. If so, that might not be enough to guarantee the truth of the counterfactual claim. A world in which I am in poverty, happy, and have a radically different attitude towards wealth than I actually have would not be sufficient to establish the truth of the sentence 'If I were in poverty, then I would be happy.'

One could avoid this problem by limiting the worlds considered to the ones that are like this world in the important respects. So, 'If I were in poverty, then I would be happy' is true just in case I am happy in the possible world that is closest to the actual world except that I am in poverty. To say that one world is close to another world is just to say that the two are similar. The world that is closest to the actual world with the exception of my being in poverty is the world that is most like the actual world except for that difference and any other difference that is required in order for that difference to be actualized.

Several extremely important philosophical concepts are often expressed as counterfactual conditionals. Some of these include law-like generalizations such as the claim that sugar is soluble. One might think that to say that sugar is soluble is simply to express the conditional that if it were dropped in water, it would dissolve. Presumably, this is true of a cube of sugar even if it is never dissolved in water. So, a cube of sugar is soluble just in case, in the possible world closest to the actual world in which it is dropped in water, it dissolves. The sugar's failing to dissolve in a world that has different physical laws than the actual world has will not be sufficient to show that the counterfactual is false.

Another concept that might be analyzed in this way is causation. To say that the striking of the match caused the match to light implies, as Hume rightly pointed out in his second definition of 'cause,' more than simply that these two events occurred. It implies that had the match not been struck, it would not have lit (Hume 1975, 76). If so, then it becomes relatively easy to provide a possible worlds analysis of that counterfactual which expresses this particular instance of causation. Unfortunately, it will not serve as a general analysis of causation. Particularly, this analysis fails in cases of preemption and overdetermination. For instance, imagine that someone strikes a match which has been laying under a magnifying glass in the sun. In this case, it appears that had the match not been struck, it still would have lit. So, some development of the counterfactual analysis of causation is needed, but even this simplistic analysis sheds some light on a particularly difficult concept.

The relation of supervenience has recently played an important role in many different areas of philosophy, particularly in the philosophy of mind. It is often used to express a certain view of the relationship between the mental and the physical. A set of properties, A (e.g. mental/psychological), is supervenient on another set B (e.g. physical) just in case there could be no change in A without a change in B, but a change in B does not require a change in A. So, the mental is supervenient on the physical if and only if there could be no change in mental properties without some change in physical properties, although there *could* be a change in physical properties without a change in mental ones.

The supervenience relation seems to capture part of the notion of the dependence of the mental on the physical while avoiding the problems associated with the identity theory. Specifically, it allows one to claim that the mental is nothing more than the physical, while maintaining that the same mental state is often realized by different physical states. This relation cannot be adequately expressed in non-modal terms. Imagine, instead, that there simply is (as opposed to 'could be') no change in the mental without a change in the physical. This claim is not strong enough to establish that the mental is dependent on the physical. As stated here, the relationship between the mental and the physical could then be purely accidental.

A possible worlds analysis of supervenience, then, would begin with the claim there are no two possible worlds where an individual differs in mental properties but not in physical properties. Interestingly, one can develop this analysis, making further fine-grained distinctions between different notions of supervenience. For instance, one might think that supervenience is a fact about individuals that actually exist, but not about all possible individuals. If so, then if an individual exists in the actual world, there are no two worlds where that individual differs in mental properties but not in physical ones. It still may be possible, then, for God to create a being whose mental states are not dependent on her physical states. Varying the analysis in this and other ways will allow one to express different conceptions of the strength or weakness of the dependency of the mental on the physical.²

One final example of the explanatory power of a possible worlds account of modality concerns the set-theoretic analysis of properties. Properties are often identified in terms of their extensions, the set of things that have those properties. For example, the extension of redness is the set of all red things. Obviously, any two properties that have different extensions must be different properties. The property of being a mammal is not the same as the property of being a dog, since the extension of mammal includes not only dogs, but other things as well.

The primary problem with extensional accounts of properties is with properties that are coextensive. The standard example is the property of having a heart and having a kidney. All things that have a heart have a kidney, but these are obviously not the same property. There are many other examples that one could use. Consider

² The claim that the mental is supervenient on the physical is often used to establish the further claim that the mental is nothing over and above the physical. One should notice that although the two claims are consistent, the first does not entail the second. At best, mental-physical supervenience entails only that there is some necessary connection between the mental and the physical.

the properties of being the first president of the United States and being the commander of the Continental Army. The extensions of these properties are identical, yet they are surely not the same property. Likewise, on this account, all properties that have empty extensions are identical. So, the property of being the King of the United States is identical to the property of being the President of Antarctica. These are all examples of properties that are coextensive but not identical.

Notice that these pairs of properties, although coextensive, are only accidentally so. It is surely possible that an object have the property of being the Commander of the Continental Army and lack the property of being the first President of the United States. George Washington, even after commanding the Continental Army, might have not been elected President; instead, John Adams might have been the first President.

So, this reasoning leads naturally to the following analysis. Two properties are identical only if it is not possible to have one but lack the other. In possible worlds terms, then, for any two properties, if there is a world in which those properties have different extensions, then those properties are not identical. Notice that necessary coextension is a necessary condition for property identity. It is not obvious that it is a sufficient condition, however. There are properties that are necessarily coextensive that are far from obviously identical. For example, the properties of triangularity and trilaterality have the same extension in every possible world, but are arguably different properties.

One will generally find this condition stated as follows: two different properties (ignoring properties that are necessarily coextensive) will always have differ-

ent extensions if one includes in the extension of the property all objects that have the property in any possible world (Lewis 1986, 51). This restatement, however, may not be equivalent to the original principle. For example, imagine that there are only three possible worlds, w_1 , w_2 , and w_3 . In world w_1 , Washington is both Commander and President. In world w_2 , Adams is both Commander and President. In world w_3 , Washington is Commander, but Adams is President. The two properties are not necessarily coextensive, since they are not coextensive in world w_3 .

If one takes the extension of a property to be simply the set containing all of the possible objects that instantiate that property, then the two properties appear to be coextensive. The extension of the property of being Commander is the set containing Washington and Adams, and the extension of the property of being the first President is also the set containing Washington and Adams. This, however, assumes that Washington in world w_1 is identical to Washington in worlds w_2 and w_3 , and likewise with Adams. If not, then the extension of the property of being Commander is the set containing Washington₁, Adams₂, and Washington₃; and the extension of the property of being the first President is the set containing Washington₁, Adams₂, and Adams₃. So, how this principle may be stated is dependent upon a particular view of identity across possible worlds, which in turn is dependent upon what kind of thing a possible world is.

These problems that have been discussed, the problem of modalities, counterfactuals, supervenience, and properties, are just some of the problems that proponents of possible worlds have addressed. It appears, then, that the concept of a possible world could have a great deal of power to explain independently important,

even vital, notions of *philosophical* concern. In order to explain anything, however, statements in terms of possible worlds must not merely be restatements of sentences containing the standard modal operators. This requires a substantive account, one that is ontologically more robust than the deflationary account. Given this apparent explanatory power, however, a substantive account is well worth pursuing.

A substantive account understands the semantics of modal logic to be providing some real insight into the nature of modality. A successful possible worlds theory of modality must, at a minimum, have three features:

(1) It must provide some insight into the nature of possible worlds.

2) It must provide an account of actuality.

3) It must include some account of trans-world identity.

The first feature requires the successful theory to provide an account of what kind of things possible worlds are. The second requires the theory to provide an account of what distinguishes the actual world from other possible worlds. The third feature is required because in order to speak meaningfully about possibilities for individuals, it appears that individuals must bear some kind of trans-world identity relation. A possible worlds theory must either explain the nature of this relation, or explain why trans-world identity is not really required.

Over the past three decades, several competing accounts of possible worlds have been proposed. These accounts can, in general, be divided into two types: modal actualism and modal realism. The proponents of both types believe that the phrase "possible world" does in fact refer to something. To say that a statement is true in at least one possible world is to commit oneself to the existence of that possible world. To the question, "Do possible worlds exist?" the actualist and realist both answer emphatically "yes". However, to the question, "What kinds of things are they?" their answers are radically different.

Modal Realism

Modal realism is a view that is most often associated with David Lewis. He introduced the view in an early work (Lewis 1968), but provided a detailed defense in his 1986 book, *On the Plurality of Worlds*. Stated simply, modal realism is the thesis that our world, the *actual world*, is simply one world among many others, and that these other worlds do not differ in kind from our world. Just as our world is inhabited by many things, including us, other worlds are inhabited by other things, some of which may be very much like us.

The parts of possible worlds are possible individuals. Any two objects that inhabit the same possible world are called "worldmates." A possible world, then, is simply the mereological sum of all of its parts. Anything that is a worldmate of part of a world is itself a part of that world. Worlds are individuated by the spatiotemporal relations of their parts. Any two things that are *spatiotemporally* related are worldmates. I and the most distant star in the galaxy are worldmates, since there is a spatial distance (a very great one, admittedly) between us. In fact, the planet that was destroyed millions of years ago when that star became a supernova is a worldmate of mine, since we are related both spatially and temporally. No part of any world is spatiotemporally related to any part of another world. Therefore, questions such as "How far are other possible worlds?" are completely misguided.

Not only are all spatiotemporally related things worldmates, but all worldmates are spatiotemporally related. The constraints on the spatiotemporal relation are fairly loose. For instance, Lewis freely admits worlds in which there is an infinite distance between worldmates. He imagines a world that is composed of temporal epochs, each epoch resembling the real line. The world, then is like many copies of the real line placed end to end. An individual existing in one epoch is a worldmate of an individual in another epoch, but there is an infinite temporal distance between them. Each of these world epochs is like a world in itself, but is not by definition a separate world. Thus, some parts of worlds might seem to be disconnected from other parts without actually being spatiotemporally disconnected. It is not a requirement that everything in the world be related in exactly the same way spatiotemporally; this allows for the possibility of spiritual beings, ghosts, and the like. Although these spatiotemporal relations are somewhat flexible, they cannot be gerrymandered in any way whatsoever. For instance, the relation of spatiotemporal disconnection will not, for obvious reasons, count as a spatiotemporal relation

One consequence of defining a possible world as a mereological sum of spatiotemporally related things is that it is then necessary that something exist rather than nothing. There are no possible worlds that have no parts; so on Lewis' account, something must exist. Does this constitute an objection to Lewis' realism? Maybe so, depending on one's intuitions. At any rate, the objection is not devastating. The claim that something must exist is compatible with the intuition that that something may not be much of anything; for example, the world might contain nothing but a singular point of unoccupied space-time. This is surely not much different from nothing at all. It is also important to note that it only follows that each world contains at least one thing. Of course, this does not entail that there is a thing that every world contains. One cannot conclude from this that there are any necessary beings. Lewis' modal realism won't suffice as a new version of the Ontological Argument.

Given the spatiotemporal disconnection between worlds, one would expect that there could be no causal relations between them either. If trans-world causation were possible, it could not be analyzed with the standard counterfactual analysis of causation. Remember that event A caused event B whenever the counterfactual, if A had not occurred, B would not have occurred, is true. In ordinary cases of intraworld causation, this counterfactual will be true at the world in which the cause and effect occurred. In cases of trans-world causation, however, there is no world in which it makes sense to claim that the counterfactual holds in any real sense. It would be true in the world where the cause occurred, but that counterfactual should be irrelevant to causation. There, had the cause not occurred, it is true that the effect would not have either; but the effect did not occur there even when the cause did. The counterfactual of causation is just false in the world where the effect occurred. The closest world to it in which the cause does not occur is that world itself, and there the effect does occur! So, there is no way to analyze trans-world causation in terms of counterfactual claims. This, however, was one of the apparent benefits of a possible worlds ontology. So, denying the possibility of trans-world causation seems to be no great loss.

How many worlds are there? The simple answer is "just enough." Lewis states that "(1) Absolutely every way that a world could possibly be is a way that some world is, and (2) absolutely every way that a part of a world could possibly be is a way that some part of some world is." This is intended to mean, presumably, that there are at least as many worlds as there are distinct possibilities. Lewis concedes that, on the realist account, these claims have very little content. Since the realist identifies ways that a world could be with possible worlds, then (1) is the trivial claim that absolutely every way some world is, is a way that some world is; or every world is identical to some world. If there are only five worlds, the first principle is still true.

So, a realist reading of principles (1) and (2) does not guarantee that there will be enough worlds to sufficiently account for all possibilities; a further principle is required. For this, Lewis develops a *principle of recombination*. The principle of recombination assumes the Humean thesis that there are no necessary connections between distinct existents. Any two things can coexist so long as they occupy distinct spatiotemporal locations. So, parts of different possible worlds are recombined to form other possible worlds. These recombined parts are not numerically identical to the parts of the original worlds; if they were, then part of one world would be spatiotemporally related to part of another world. Instead, these parts must be identical in the sense of being a duplicate. There cannot now simply be five worlds. The number of worlds is very large, presumably uncountably so (Lewis 1986, 86-90).

What makes our world, the actual world, different from the others? The intuitive response is that the actual world exists and the others do not, but this response is not available to the realist. Quantification over worlds requires that all worlds exist, and for the realist, they exist in exactly the same manner. So instead, Lewis offers what he calls an "indexical analysis" of actuality. 'Actual' is an indexical term, like 'here' and 'present.' The reference of indexical terms varies according to the context in which they are used. 'Here' refers to the place at which the statement containing the term was uttered; 'present' refers to the time at which the statement was uttered. So, 'actual' refers to the world at which the statement was uttered. So, 'actual' refers to the world at which the statement was uttered. To say that there are no actual leprechauns is true, even though there may be (and presumably are) leprechauns that exist in other worlds. To say that there are no actual leprechauns in the world in which that sentence is uttered. Does this mean that all worlds are actual? No, this world is the only actual world; just as this time is the only present time.

In possible world terms, to say that Bob Dole might have won the election in 1996 is to say that there is a possible world in which Bob Dole did win the election in 1996. At first glance, it must be the case that the name 'Bob Dole' refer to the same person in each possible world. When I say that Bob Dole might have won the election, I don't intend that to be true just in case someone named Bob Dole won in some possible world. Quantification over individuals in worlds seems to require that the same individual exist in different worlds. So, a theory of possible worlds, if it is to serve as an analysis of modal claims, must provide some account of trans-world identity.

It should be obvious that the realist must deny trans-world identity. Since worlds are disconnected, no two worlds share any of the same parts. If a being were to exist both in the actual world and some other possible world, then the same individual would always occupy two different spatiotemporal locations. So, Bob Dole, being part of this world, exists in no other world. Therefore, on the realist account, it seems that it is not possible for Bob Dole to have won the election.

In order to avoid this implication, Lewis introduces the *counterpart relation*. The name 'Bob Dole' refers to Bob Dole in the actual world, and to the counterpart of Bob Dole in every world in which he has a counterpart. So, the statement that Bob Dole might have won is true if and only if there is a possible world in which a counterpart of Bob Dole won. The counterpart relation is a similarity relation; not an identity relation, not even an equivalence relation in most cases. The counterpart of Bob Dole in possible world w resembles the actual Bob Dole more than anything else does in possible world w.

The counterpart relation is reflexive, non-symmetric, and non-transitive. Imagine three possible worlds, w_1 , w_2 , and w_3 . Stipulate that w_1 is the actual world, which contains among its inhabitants Lassie, a collie. All the dogs that exist in w_2 are beagles, but w_3 includes both beagles and collies. The counterpart relation must be reflexive. If it were not, then it could be the case that something in the actual world resembles the actual Lassie more than Lassie herself does, but this is clearly impossible. The relation is non-symmetric since the counterpart of Lassie in w_2 must be a beagle, but the counterpart of that dog in the actual world will not be Lassie, it will instead be an actual beagle. Likewise, the counterpart relation is nontransitive. Lassie's counterpart in w_2 is a beagle, and that dog's counterpart in w_3 is also a beagle, but Lassie's counterpart in w_3 must be a collie instead. Something might have more than one counterpart in a particular world. Imagine that a certain pair of identical twins inhabit some possible world, and these twins resemble the actual Bob Dole more than anything else does in that world. If no twin resembles him more than the other twin does, then Bob Dole has two counterparts in that world. Of course, there must be worlds in which Bob Dole has no counterpart. If there were no such worlds, then the sentence "Bob Dole exists" would be true in every possible world, elevating Dole to the status of a necessary being. (Lewis 1986, 11) Presumably, there are worlds in which nothing exists in those worlds that resembles Dole much at all (Lewis 1986, 112-113).

Robert Stalnaker summarizes Lewis' realism in four theses:

- (1) Possible worlds exist. They are just as real as the actual world.
- (2) Other possible worlds do not differ in kind from the actual world.
- (3) 'Actual' is an indexical term, referring to the world in which it is uttered.
- (4) Possible worlds can be reduced to nothing more basic (Stalnaker 1976, 67-68).
- To these four theses, I add:
- (5) Trans-world identity is analyzed in terms of similarity between counterparts.

At this point, I have discussed all of these theses except for (4), the claim that possible worlds cannot be reduced to some more basic entity. Lewis' reasons for holding this thesis are found in his objections to alternative views of possible worlds. So, before discussing his reasons for this thesis, it may be helpful to examine some of those views.

Modal Actualism

Most proponents of possible worlds have argued that we can enjoy the benefits of possible worlds theory without paying the ontological cost of Lewis' modal realism. They have argued that the framework that provides these benefits can be developed with an ontology that does not allow quantification over nonactual objects. The only objects that exist are those that actually exist, hence the name *modal actualism*. There are many varieties of modal actualism, and proponents of actualism disagree with each other in significant ways. Three of the most prominent versions of modal actualism are those proposed by Robert Stalnaker, Robert M. Adams, and Alvin Plantinga. On their accounts, merely possible worlds are either identified with things that actually exist, or are constructed out of things that actually exist. These accounts differ as to which actual objects will best serve as possible worlds. Here, I will examine ways that actualists have approached the three issues concerning possible worlds theories: the nature of possible worlds, the nature of actuality, and the nature of trans-world identity.

Robert Stalnaker was one of the first to address Lewis' realism. In his paper "Possible Worlds", he offers what he calls "a more moderate form of realism." Stalnaker accepts the first thesis of modal realism, that other possible worlds exist. He is careful to point out, however, that this thesis says nothing about the nature of these worlds. It is the second thesis that explicitly does the ontological work. This is the claim that other possible worlds are of the same kind as the actual world, which is simply described as "I and my surroundings." (Stalnaker 1976, 67)

Stalnaker finds Lewis' paraphrase argument plausible, that possible worlds are simply ways the world could have been. The important question, then, concerns the nature of these "ways the world could have been," specifically, are they entities of the same kind as the actual world? This, Lewis' second thesis, is what all three actualists discussed here deny.

Stalnaker argues that the reasons for accepting the first thesis, the paraphrase argument for possible worlds, is in fact incompatible with the second thesis. The second thesis has two parts. The first part is a claim about the relation that other worlds bear to the actual world: they are of the same kind. The second part is a claim about the nature of the actual world: it is simply I and my surroundings. Now, Stalnaker points out, if Lewis really believed that other possible worlds are ways things might have been, he ought to deny the second part. He ought to instead claim that the actual world is the way things are, which is a plausible claim indeed.

What does it mean to say that the actual world is the way things are? This depends on the use of 'is' in this sentence. Is it a statement of identity or one of predication? If it is the former, then exchanging the position of the terms should not result in an awkward reading of the sentence. The statements 'Mark Twain is Samuel Clemens' and 'Samuel Clemens is Mark Twain' are equally felicitous. Statements of predication, such as 'The film is dull' are generally awkward (if not simply ungrammatical) when reversed. The statement 'Dull is the film' has a poetic ring to it, but imagine hearing every English predicative in this form. Stalnaker affirms that "The world is the way it is," and this certainly seems to be true. Revers-

ing that, one gets "The way the world is is the world," which is certainly very awkward.

He concludes that 'the way things are' does not name something that is identical to the world, but names a *property* that the world instantiates. 'Ways things might have been' must then name properties that the world might have instantiated but did not. Other possible worlds are uninstantiated world properties. If properties can exist uninstantiated, then it is possible that world-properties exist in the actual world, although only one of them is instantiated (Stalnaker 1976, 68).

Arguments that draw metaphysical conclusions from natural language usage are notoriously less than compelling. Especially in this case, for Stalnaker has begged the question in his statement, "The world is the way it is." If he were consistent with his earlier statement, then he should have phrased it "The world is the way things are." The corresponding claim would then be "The way things are is the world," which I, in any case, do not find awkward. So, I do not find this natural language argument for possible worlds as properties compelling.

Another objection is that Stalnaker's actualism requires not only that uninstantiated properties exist, but also that they exist necessarily. If a particular worldproperty could fail to exist, then it could have been the case that that world not be possible. If it is possible that a world not be possible, then in S-5, the world is simply not possible. Unfortunately, Stalnaker's claim that properties *could* exist uninstantiated is not strong enough to satisfy this demand. Showing that properties exist necessarily requires argument, and in the absence of further argument, acceptance of Stalnaker's account of the nature of possible worlds depends entirely on the availability of plausible alternatives.

Fortunately, the actualist is not limited to properties as the only candidates for possible worlds. In fact, properties are not even the most obvious choice. In the *Tractatus*, Wittgenstein writes "The specification of all true elementary propositions describes the world completely...It is possible for all combinations of atomic facts to exist, and the others not to exist." (Wittgenstein 1922, 4.26-27) Building upon this, one might develop an account of possible worlds as sets of propositions.

Jaako Hintikka was one of the first to systematically develop such a view (Hintikka 1979, 153-157). Later, Robert M. Adams developed a similar account in his work on actualism. For Adams, a possible world is a "world-story," which is a maximally consistent set of propositions. A maximal set is a one such that, for any proposition, it contains either that proposition or its negation. It is consistent just in case it is possible for all of its members to be true. Each world story is a complete description of a way the world could have been. It must be complete, since adding to it any proposition that it does not already contain will guarantee its inconsistency. Since it is maximal, if it does not contain some proposition, it must already contain the negation of that proposition. The world story that contains all true propositions is the complete description of the actual world (Adams 1974, 225-228).

So, Adams affirms the first thesis of modal realism, that possible worlds exist. He denies the second thesis, that other possible worlds do not differ in kind from the actual world. In fact, every actualist theory must deny this thesis. For the actualist, only actual objects exist; everything that exists is an actual object. If the actual world is the totality of all that exists, then other possible worlds are proper parts of the actual world. Therefore, other possible worlds cannot be the same in kind as the actual world.

A closely related view is that of Alvin Plantinga. Plantinga's position seems to capture the plausible intuition behind the paraphrase argument of David Lewis: that possible worlds are the way things could be, and the actual world is the way things are. The proposition 'Philadelphia is the capital of the United States' is not a way things could be, rather, Philadelphia's being the capital is a way that things could be. Philadelphia's being the capital is a state of affairs. In fact, it is one that fails to obtain. Other states of affairs do obtain, such as Washington D.C.'s being the capital of the United States. For Plantinga, possible worlds are reduced to states of affairs.

States of affairs can both include and preclude other states of affairs. For example, Washington D.C.'s being the capital of the United States includes the United States' having a capital, but precludes Philadelphia's being the capital. Obviously, not just any state of affairs can serve as a possible world. Like Adams, Plantinga requires that these states of affairs be maximally possible. S is a maximal state of affairs if for every other state of affairs S^* , S either includes S^* or precludes S^* . S is a possible state of affairs just in case it is possible for it to obtain.

The actual world is the maximally possible state of affairs that does in fact obtain. There are other maximally possible states of affairs that do not obtain. Since the actual world obtains, at least one world does obtain. In contrast to Lewis, at most one world can obtain. If two distinct worlds obtained, then there would be some state of affairs included by the first and precluded by the second. In that case, the state of affairs both would and would not obtain, which Plantinga finds impossible.

Sets of propositions, however, still play an important role in Plantinga's account. There is a clear relation between the University of Oklahoma's being located in Norman and the proposition expressed by the sentence 'The University of Oklahoma is located in Norman.' The state of affairs obtains if and only if the corresponding proposition is true. This correspondence will clearly hold if states of affairs are identical to propositions as Chisholm maintains (Chisholm 1970). If so, then at this point, Plantinga offers nothing more than Adams. If not, then there are certain differences, but for now these differences are quite minimal. For in any case, every maximally possible state of affairs corresponds to exactly one maximally consistent set of propositions, which Plantinga calls a "world book." (Plantinga 1974, 44-46) So, for both Plantinga and Adams, possible worlds are reduced to something more basic. A possible world is either a maximally consistent set of propositions, which is constituted by its elements; or it is a maximally consistent state of affairs, constituted by simpler states of affairs.

The second issue facing a theory of possible worlds is the nature of actuality. On the surface, it appears that the actualist would have a particular problem here. For the actualist, everything that exists actually exists. If possible worlds exist, then they must actually exist. What sense, then, does it make to speak of *the* actual world? Of course, this apparent problem is not likely to trouble the actualist. As has been shown, actualism requires that merely possible worlds be proper parts of the actual world, so they cannot be on a par with the actual world. Even though merely possible worlds actually exist, on all three versions of actualism, there is a fundamental difference between them and the actual world.³

The primary question that faces the actualist concerning this issue is whether to accept Lewis' indexical analysis of actuality. Stalnaker explicitly affirms the thesis that the indexical analysis is correct. Plantinga does not explicitly affirm the indexical analysis, but his discussion of actuality in a world at least suggests such a view. This acceptance of the indexical analysis seems to present a problem. If inhabitants of other worlds can truly affirm the actuality of their world, then there seems to be no real difference between the actual world and these other worlds. Actuality, then, is nothing particularly special, despite the insistence of the actualists.

According to Plantinga, each possible world is actual in itself. This is to say nothing more than the proposition 'This world is the actual world' is part of every world book. Here, 'this' is used indexically, referring to the world for that particular world book.

This does not entail that all worlds are actual, however. To say that something is possibly true is just to say that it is true at some world. For the actualist, however, to say that a proposition is true at some world is simply to say that if that world were actual, then that proposition would be true. So, the sentence simply means that if this world were the actual world, then this world would be the actual world. Since it is a tautology, it must be true in every possible world; a world that

 $^{^3}$ The phrase 'merely possible' is used to distinguish other possible worlds from the one possible world that is the actual world.

fails to be actual in itself, is one that could not be the actual world, and therefore is not a possible world (Plantinga 1974, 48-49).

It is important to notice that the indexical analysis is a *semantic thesis*; it is purely a claim about the meaning of 'actual.' This semantic thesis is distinct from the metaphysical claim that there is more than one world. In fact, one can affirm the indexical analysis of actuality, and maintain that there is only one world, just as one can hold an indexical analysis of 'I' and yet be a solipsist. So, the indexical analysis does not commit one to any specific views of the nature or number of other worlds (Stalnaker 1976, 69).

Adams, however, explicitly rejects the indexical analysis. Actuality, for Adams, is not relative but absolute. The actuality of a possible world consists in the propositions contained in that world-story being true, not true in that world, but true in fact. He presents three arguments for his rejection of the indexical analysis. First, if the indexical analysis is true, there is no actual world in the absolute sense. A world is actual only in relation to its inhabitants. Actuality, then, conveys no special status upon a world except from the perspective of its inhabitants. If one thinks that actuality is a special status, then one should deny the indexical analysis.

This argument fails, however. The indexical analysis of actuality does not require that every world be actual in the same sense as the actual world. It only requires that every world affirm its own actuality. In the actual world, Roger Bannister was the first to run a mile in less than four minutes. Other possible worlds will affirm that someone else was actually the first to do so. What makes Bannister special is not that a world affirms that he was actually the first, but that he was actually the first. Affirming one's actuality may be nothing special, but being actual certainly is.

Adams' second argument is that the indexical analysis conflicts with our ordinary value judgments. It is not the possibility of a vicious action that is bad per se, but its actuality. It is not simply the possibility of the Holocaust that is so disturbing, but the fact it actually occurred. Adams maintains, however, that on the indexical analysis, even the possibility of the Holocaust is enough to guarantee that it is actual at some world. So, we should find the possibility of evil as bad as the actual evil itself. If so, then it seems as if we have no reason to refrain from actualizing some evil, since it will be actualized at some other place anyway. On this view, preventing a possible evil in the actual world is the same as ensuring that it occurs in some other world.

This argument also fails. First, even if possible evils are actualized in other worlds in the same sense that things are actualized in the actual world, we may still have good reason to prevent their occurring in the actual world. One might grant that even if the overall balance of good and evil cannot be changed, one still has the obligation to prevent specific acts evil when one can. This argument is analogous to reasoning that it makes no difference if I steal someone's wallet, because if I don't, someone else surely will. It is also plausible to think that there is nothing wrong with being more concerned with evils that are closer to us than those that are farther away (Beedle 1996, 490),

Adams' third argument requires closer examination. This argument concludes that the indexical analysis conflicts with judgments of identity in different

possible worlds. Adams proposes a sufficient condition for trans-world identity of events and individuals. Since the argument concerns the identity of events, I will focus on that principle:

 C_e : If an event *e* occurs in possible world *w* at time *t*, and event *e** occurs in possible world *w** at time *t**; and (2) the whole history of *w** up to and including *t** (and no other time in *w**) is precisely the same, qualitatively, as the whole history of *w* up to and including *t* (and no other time in *w*); and (3) the whole previous history and present state of *a** (and of no other individual in *w** at *t** is precisely the same, qualitatively, as the whole previous history and present state of *a** (and of no other individual in *w** at *t** is precisely the same, qualitatively, as the whole previous history and present state of *a* (and of no other individual in *w*) at *t*; then *a* is numerically identical with *a** (This is Adams' principle C, modified as he suggests in Adams 1974, 217-218.)

Now, imagine writing two sentences: (1) "Actually, men will land on Mars by 2100 A.D." and (2) "Actually, it is not the case that men will land on Mars by 2100 A.D." It is plausible to assume that there are worlds in which humans do land on Mars by that date, and worlds in which humans do not. It is also plausible to assume that at least one of both types of world has precisely the same history as the actual world up to, and including, the time of your writing the sentences. So, your act of writing (1) occurs in probably many different worlds, and it is true in some of those worlds and false in others.

Adams discusses several possible meanings of 'actually' in these sentences before turning to the meaning that the person holding the indexical analysis probably prefers: 'actually' means 'in this possible world' and unambiguously designates the world in which the sentences were written and no other. The problem is that, given C_e , the events are identical. How can the same event unambiguously designate more than one thing?

There are at least two responses that can be made to this argument. First, one could simply deny C_e , which a counterpart-theorist would surely do. An actual-

ist is not likely to do so; anyone who is committed to the trans-world identity of individuals should also be committed to the trans-world identity of events. We speak meaningfully about events in different possible worlds, such as "Bill Clinton's conviction could have led to an impeachment." This sentence seems to require the trans-world identity of both Clinton and some event, namely, an impeachment.

It may be that there is a need to deny C_e , but it is not because of a conflict with our intuitions concerning actuality. If this argument is successful, then it is successful even without the adverb 'actually.' For example, consider the sentence "Men will land on Mars by 2100 A.D." This sentence is surely true in some worlds but false in others. So, any prediction I make will result in truth at some world or another. If this argument shows that an indexical use of actually makes predictions impossible, then it shows that all predictions are impossible.

One can stipulate that any usage of 'actually' that obtains unambiguously designates the world that obtains. Adams' third argument suffers from a flaw that also plagues the second argument: Adams assumes that the indexical theory entails that more than one world can be actual in an absolute sense. This assumption, however, is false. The indexical theory alone makes no assumptions about the number of actual worlds in an absolute sense, it is the indexical theory plus a conception of the nature of possible worlds. For actualists such as Stalnaker and Plantinga, the preferred conception of possible worlds guarantees that only one can be the actual world.

The third issue concerning a modal theory is *trans-world identity*. Intuitively, trans-world identity is required for a possible worlds analysis. If it was possible that Bob Dole win the election, then there is a possible world where Bob Dole won the election. Surely, however, the name 'Bob Dole' must refer to the same person to which it refers in the actual world. How could someone else's winning the election establish the claim that Bob Dole could have done so?

As we saw earlier, the modal realist cannot affirm trans-world identity in a strict sense, because all worlds are causally disconnected with no parts in common. So, the modal realist must affirm counterpart theory. The need for disconnected worlds arises only because of the realist's account of the nature of worlds. Since the actualist denies such an account, the actualist may be free to affirm a robust conception of trans-world identity. It is far from obvious, however, that such a conception is tenable.

Roderick Chisholm (1963) was the first to raise doubts concerning the notion of trans-world identity. In a later article, he argued that trans-world identity leads to situations that violate both the indiscernibility of identicals and the identity of indiscernibles.

Chisholm asks us to consider the person Adam who, in the Biblical account, lived to the rather mature age of 930. Someone who lived to be 930 surely could have lived to be 931, after all, that is only a 0.1% increase in his life span. So, transworld identity requires that there be a world w_1 in which Adam lived to be 930, and another world w_2 in which the same individual lived to be 931. They are identical, but not indiscernible. This, as Chisholm points out, is not enough to show that trans-world identity is fraught with problems. In this example, identity across

worlds is no more problematic than identity across time, and most of us are not prepared to deny trans-temporal identity.

The problem quickly escalates. In w_1 Adam lived to be 930, and Noah lived to be 950. Now, in w_2 , Adam lived to be 931, and Noah lived to be 949; and in w_3 , Adam lived to be 932, and Noah lived to be 948. Imagine a continuing succession of worlds, $w_4...w_{2l}$, in which Adam lives one year longer than in the previous world, and Noah dies one year earlier. In world w_{2l} , Adam lived to be the same age as Noah lived in w_1 and vice versa. Now, imagine other successions of worlds in which Adam and Noah trade other properties. Eventually, there will be a world w_n in which Adam in w_n and Noah in w_l share exactly the same properties, as do Noah in w_n and Adam in w_l . Now, it appears that there are two beings that are qualitatively identical but not numerically identical (Chisholm 1979, 80-84).⁴

According to Chisholm, the only possible solution is in terms of the essential properties of Adam. The notion of an essential property goes back at least as far as Plato and Aristotle. In contemporary terms, a property F is essential to an object x, just in case x has F in every world where x exists. So, it is impossible for x to exist and lack F.

The notion of an essential property, in itself, is not enough to solve the problem of trans-world identity. Having that particular essential property is a necessary condition for an object's identity, but the problem of trans-world identity requires a sufficient condition for being that very individual object. For example, if existence

 $^{^4}$ For the moment, I will assume that it is possible for two numerically distinct things to be qualitatively identical. This issue, and the notion of essential properties in general, is examined in detail in Ch. 5.

is a property, then it is clearly an essential property; it is impossible for Adam to exist and lack existence. Existing, however, will not guarantee that some object in a world is identical to Adam. The problem of trans-world identity requires that there be a property, whether simple or compound, that guarantees identity. If such a property exists, then it is impossible for anything to have the property and not be Adam; likewise, it is impossible for anything to be Adam and lack the property. If *A* is the property of being Adam, then $\Box \forall x(Ax \equiv x=a)$. Properties such as this are often called "haecceities" (a term coined by Duns Scotus) or "individual essences".

The notion of an individual essence is used by Plantinga to solve the problem of trans-world identity. Two things are identical if and only if they exemplify the same individual essence. Whatever other properties may change, as long as the properties that constitute the individual essence of Adam do not change, then Adam retains his identity (Plantinga 1979, 262-268).

In Chisholm's example, if Adam in w_l is identical to Adam in w_n , it must be in virtue of Adam's individual essence. This cannot be a qualitative essence, since they share no qualitative features. Identity, if it is maintained in this case, must be in virtue of a non-qualitative property that is shared by the two: the property of being Adam. Robert Adams calls such a property a "primitive thisness." (Adams 1981)

The notion of an individual essence not only allows the actualist to explain trans-world identity, but to maintain consistency with our intuitions in a few special cases. One particular case concerns the possibility of objects that do not actually exist. It is obvious why this should present a difficulty for the actualist. It does seem to be the case that there could have been something that does not exist in the actual world. On a standard possible worlds analysis, this requires that there be an object that exists in some world but not in the actual world. According to actualism, however, there are no objects in any world that do not exist in the actual world.

Plantinga's solution is to appeal to the (alleged) necessary existence of properties. Individual essences, being properties, exist necessarily. Their existence in every possible world does not entail that they are also exemplified in every possible world. If Socrates is a contingent being, then there are worlds where Socrates does not exist. In those worlds, the individual essence of Socrates exists, it just fails to be exemplified. So, there are possible objects that do not actually exist just in case there are individual essences that are not exemplified in the actual world (Plantinga 1979, 268-272).

It appears, then, that there are ways that an actualist can meet the fundamental requirements of a modal theory. This does not mean that modal actualism is automatically preferable to modal realism. Any version of modal actualism will be committed to something that some philosophers find troublesome, such as properties, propositions, or sets. Of course, there are other reasons to have such things in one's ontology, so they provide little reason for rejecting actualism. Other notions, however, such as haecceities and primitive thisnesses are likely to be much more problematic. There are also many ways besides ontological simplicity in which one theory may be preferable to another, and it might be the case that the ontological cost of realism is worth the price. These are the issues for the next section.

Evaluating the Two Accounts: Realism vs. Actualism

Obviously, only the actual world can be observed (at least by actual people). A theory about the nature of possible worlds, then, is a theory about something that is in principle unobservable. The theories will have to be evaluated on nonempirical principles. For this reason, some are tempted to dismiss the entire inquiry as useless, but this is not just a problem for speculative metaphysics. Since the available evidence always underdetermines the theory, we have been forced to develop several principles to use for non-empirical theory confirmation. The most important of these principles are internal consistency, initial plausibility, consistency with other theories that are accepted as true, theoretical simplicity, and explanatory power.

First of all, a successful theory must be *consistent*. If the theory is regarded as true, then anything that follows deductively from the theory must also be regarded as true; but if the theory is inconsistent, any statement and its negation may be derived from it. Surely, both cannot be regarded as true. Since any statement whatsoever can be derived from an inconsistent set, an inconsistent theory will lead to falsehoods as easily as to truth.

If two competing theories are both internally consistent, then the other three principles are used to determine their relative worth. The *initial plausibility* of a theory is probably the least important of these three principles. The more plausible of two theories has the initial advantage, but as evidence mounts, and as the explanatory strength of each theory is evaluated, a theory with a low initial plausibility can become more and more plausible. Plausibility is used primarily when making judgments concerning which experiments to do first. A good theory should also be consistent with other theories that have been strongly confirmed. Of course, a theory that conflicts greatly with statements that are currently accepted as true will have a very low initial plausibility. If there is such conflict, it may be the case that the previously accepted theories should be dismissed, but in any case, the fewer revisions to the network of belief required, the better off the theory is.

The next principle is *theoretical simplicity*. Simpler theories are preferable to complex theories. Simpler theories are easier to use when making predictions, and, to some extent, aesthetically pleasing. Most importantly, one can make the case that simpler theories are more likely to be true (see Quine 1976b). Of course, the notion of simplicity is itself not all that simple; it is not at all obvious what should count as the proper parameters of simplicity. A theory can be simple in that it makes very few axiomatic assumptions, or has only integers for constants. The most commonly proposed type of theoretical simplicity, however, is ontological parsimony: a successful theory should make as few ontological commitments as possible (Quine 1976a).

The final principle is *explanatory power*. A good theory is one that helps us understand why or how something is the way it is. The deeper the level of understanding and the wider the scope of understanding, the more successful the theory. The classic example is Boyle's Law and the Kinetic Theory of Gases. Boyle's Law states that at a constant temperature, the volume of a confined ideal gas varies inversely with its pressure. The Kinetic Theory of Gases states that gases are made up of extremely small particles which are constantly in motion. The pressure of the gas
is determined by the collision of the particles with the container, and the temperature is determined by the average energy of the particles. Obviously, the explanatory power of the Kinetic Theory is much greater than that of Boyle's Law. The Kinetic Theory not only explains anything that might be explained by Boyle's Law, but, if true, it also explains why Boyle's Law holds.

An important thing to note is that, with the exception of the requirement for internal consistency, these principles are *ceteris paribus* principles. An internally consistent theory is always preferred to one that is inconsistent. With respect to the other requirements, however, a theory might win in some instances and lose in others. All other things being equal, a simple theory is preferred to a complex one. It is not the case, however, that a simple theory will always be preferred to a complex one. A complex theory with a great deal of explanatory power might be preferred to a simple theory with little explanatory power. Now the task is to determine how these principles apply to the modal theories that have been proposed.

Initial Plausibility

For the moment, I will assume that both realism and actualism are internally consistent. The initial plausibility of a theory is a subjective notion, such judgments will vary from person to person. Given the standard reaction, however, I gather that most would judge modal realism to be less plausible than actualism. Early, David Lewis claimed to have met with few arguments against his position, but many incredulous stares. I take incredulous stares to be good evidence against initial plausibility (Lewis 1973, 86). Of course, a theory with a very low initial plausibility can, in the end, be the successful theory. I am sure that the initial plausibility of reductive materialism and quantum mechanics were both quite low, so a low initial plausibility is no reason to automatically reject realism. We must instead examine how it fares with the remaining principles. (Bigelow and Pargetter 1987, 100-102)

In this case, it is difficult to apply the principle that requires a good theory to be consistent with well-confirmed theories, since, with respect to modalities, there appear to be no other theories with which to be in conflict. Nevertheless, we want the best theory to be one that is generally consistent with our pre-theoretic modal intuitions. This, however, seems to introduce a potential circularity problem. One purpose that a modal theory would presumably serve is to enable us to determine which of our pre-theoretic modal intuitions were correct. If so, then the circle of justification seems fairly small. The theory is justified by an appeal to intuitions, but the intuitions are justified by an appeal to the theory.

The danger is not as great as it seems. In this respect, a theory of modality is quite similar to a theory of ethics. Ethical theories are often judged by an appeal to pre-theoretic moral intuitions, then used to determine whether certain moral intuitions are also justified. Circularity is avoided, however, because the intuitions that justify the theory are not the same as the ones that are justified by the theory. For example, a moral dilemma is a situation in which I might have intuitions about what I should do, but I am not at all sure that those intuitions are correct. I need a theory to help me determine how I should act in that particular situation. There are many moral theories from which I can choose, however, and I need some reasons to rule some out in favor of others. Not all of my intuitions concern problematic situations; some of them are very clear and unquestioned. For instance, my intuitions are that child abuse, rape, and torture are all wrong, and I will reject any moral theory that conflicts with those intuitions. So, I have many intuitions concerning possibility and necessity; some of them are questionable, but others are not. For example, any modal theory that denies the possibility of the actual is unacceptable, as is any theory that affirms the possibility of logical contradictions.

Both types of theories meet these conditions. There are some less trivial intuitions which some have claimed conflict with modal realism. The objections that are commonly presented against modal realism are misguided. For example, these include the objections to the indexical analysis of actuality given by Adams which have already been discussed. A recent argument from Andrew Beedle is that realism implies fatalism:

- 1. If the character of our world is given, then our actions and choices are completely futile.
- 2. Realism holds that the totality of logical space, which is the character of all the possible worlds is given.
- 3. This implies that the character of each world, including our world, is given.
- C. So, our actions and choices are completely futile (Beedle 1996, 492-495).

Realists need not grant that the second premise implies the third. The second premise states only that the totality of possible worlds encompasses all logical possibilities. This does not in itself determine which logical possibility must be enacted by this world. Can I choose to forego another cup of coffee this evening? Ignoring issues surrounding the debate about free will and determinism, I will assume that I can. If I choose to forego another cup of coffee, the first premise only entails that I have a counterpart in some world that will not choose to do so.

The counterpart relation is a similarity relation, not a relation of strict identity. So, the other-worldly counterpart of a person is never identical with the actual person. Saul Kripke points out an apparent conflict between this and our feelings of regret concerning what might have been:

Thus if we say 'Humphrey might have won the election (if only he had done such-and-such), we are not talking about something that might have happened to *Humphrey* but to someone else, a 'counterpart'.' Probably, however, Humphrey could not care less whether someone else, no matter how much resembling him, would have been victorious in another possible world (Kripke 1980, 45 n. 13).

The objection misses the point that the counterpart theorist makes. Let us grant that in some sense of 'might', Humphrey might have won. Given that he did not actually win, but could have won, the modal realist claims that Humphrey's counterpart, a person very much like Humphrey, did win in another possible world. Kripke correctly points out that Humphrey could not care less who might have won in another possible world, even if it were someone very similar to himself. The counterpart theorist doesn't require Humphrey to care about someone else that might have won, instead the realist grants that Humphrey should care that he might have won. The important thing to Humphrey about his other-worldly counterpart is not that the counterpart might have won, but that the counterpart did win. The counterpart's winning the election in another possible world shows that Humphrey could have won in this one.

These are examples of ways that modal realism is said to conflict with our intuitions. Of course, there are those that claim that actualism conflicts with others of our intuitions. In both cases, these conflicts often rest upon a misunderstanding of the position in question. In some instances, it is not at all clear that there is even a conflict (for example, see Tomberlin and McGuinness 1994).

In the end, however, realism is the view that is least consistent with some of our most important modal intuitions. The most obvious one is the intuition that the world is the totality of existing things. If realism is true, then this intuition is just false. Of course, many of our intuitions in the past have been false. We cannot reject realism solely on this basis.

Simplicity

With respect to ontological simplicity, the theory that is committed to the fewest kinds of entities is the preferred theory. In this case, a theory that adds a new kind of entity to our ontology, a possible world, is more complex than a theory that can construct possible worlds from already existing things. So, it seems that the principle of ontological simplicity favors actualism over realism, especially if one is already committed to the existence of propositions, properties, and the like.

One shouldn't be so quick to dismiss modal realism as the less parsimonious theory, however. There are two responses available to the realist that must first be examined. The first response grants that one adds a new irreducible kind, a possible world, to the ontology. But if adding this new kind to the ontology enables us to trim several other kinds out of it, then modal realism does not result in less simplicity, but results in more overall ontological simplicity. Two things are commonly proposed as reducible to possible worlds: properties and propositions. Earlier in this chapter, I discussed the problem with a traditional extensional analysis of properties, namely that it requires coextensional properties to be equivalent. Adding possible objects to the extensions of properties allows one to circumvent this problem in most cases, although it certainly does not do so completely. A modal extensional analysis of properties will still require that necessarily coextensive properties are equivalent. So, if Goldbach's conjecture is true, then the modal extensional analysis is committed to the equivalence of the property of being an even number and the property of being the sum of two primes. If one maintains that these are different properties, then adding possible worlds to the ontology will not result in a net gain of simplicity.

Another proposed candidate for reduction is the proposition. Lewis and Stalnaker both maintain that propositions can be reduced to sets of worlds. For both, propositions are identified with properties that are instantiated by entire worlds. The proposition then, becomes the set of all worlds in which that proposition is true (Lewis 1986, 53-55; Stalnaker 1976, 72-73). The intuition behind these extensional accounts is that, when one adds possibilia to the extension, the only thing the members of the set will have in common is the thing in question. So, when one considers the set of all worlds in which some proposition p is true, the only thing those worlds will have in common is the truth of p. As in the extensional analysis of properties, however, the extensional analysis of propositions fails with respect to necessarily equivalent propositions. On this analysis, all necessarily true propositions are the same proposition. So, every mathematical truth is identical to every other mathematical truth. If one denies this, then one must deny the reduction of propositions to sets of worlds.

The second response available to the realist grants that modal realism is less parsimonious in one sense, but it is not the sense of parsimonious that is important for theory confirmation. David Lewis distinguishes between two kinds of parsimony: qualitative and quantitative. A theory is qualitatively parsimonious when it is committed to the least number of kinds of entities. A theory that is committed to sets and numbers as both basic ontological kinds is less qualitatively parsimonious than a theory that takes sets as basic and constructs numbers out of those sets.⁵ A theory that postulates seven fundamental kinds of particles is less qualitatively parsimonious than one that postulates five.

On the other hand, a theory is quantitatively parsimonious when it is committed to the least number of instances of the kinds of entities that it postulates. A theory that postulates 10⁴⁶ electrons is less quantitatively parsimonious than one that postulates 10⁴⁵ electrons. It is qualitative parsimony that is considered important in theory confirmation, not quantitative parsimony. Modal realism, Lewis maintains, only fails with respect to quantitative parsimony. It is at least as qualitatively parsimonious as actualism. Modal realism does not ask that anyone believe in any more kinds of things than they already believe. We all believe in the actual world, realists only believe in more instances of that kind (Lewis 1973, 87).

⁵ One method is to define zero as the null set, then each integer is defined as the set containing all of the previous integers. So, $0 = \emptyset$; $1 = \{\emptyset\}$; $2 = \{\emptyset, \{\emptyset\}\}$; and so on.

This response may be convincing so long as one does not give it much thought. Even if one grants that realism and actualism are each equally qualitatively parsimonious, and that qualitative parsimony is the type that is most important for theory confirmation, it should not be automatically conceded that quantitative parsimony plays absolutely no role in theory choice. Surely, there is a significant difference between a theory that postulates one instance of an entity, the world, and another that postulates an infinite number (more than likely, an uncountably infinite number) of instances of that kind.

It is not the case, though, that Lewis simply asks us to believe in more instances of a kind to which we are already committed. In fact, he asks us to believe in many more kinds. Is it possible that the world contain one more kind of particle than it does? If so, then some possible world does so. Is it possible that unicorns exist? If so, then they do in some possible world. Modal realism fails the test of ontological simplicity, no matter what sense of simplicity one has in mind.

Actualism does not itself completely avoid postulating questionable entities. Two problems traditionally plague the actualist. The first problem is explaining how identity is maintained across worlds between individuals that differ qualitatively. The second is to explain meaningful statements concerning possible but nonactual individuals. Plantinga solved both of these problems by appealing to individual essences. If these essences are qualitative, then there is no real ontological cost. This is a case, however, where the old adage rings true: "You get what you pay for." Could there be two qualitatively identical individuals? If so, assuming that essences are qualitative, then either these individuals have the same individual essence, or they have no individual essence. It is impossible that two distinct individuals have the same individual essence, so there are individuals that have no essence. The problem, then, is that the actualist cannot speak meaningfully about such possible individuals. In order to do so, she must refer to uninstantiated essences, but there are none for these individuals to possibly instantiate.

So, the actualist must be committed to non-qualitative essences. Such essences cannot be reduced to general qualitative properties. Therefore, they must always somehow contain the proper name of the individual, e.g. 'the property of being Socrates.' In order to solve the problem of possible but non-actual individuals, these essences must refer by name to individuals that do not actually exist, but how can such an essence exist in a world in which the individual does not also exist? To say that such things exist necessarily is, as Alan McMichael points out, "too glib." (McMichael 1983, 56-61) So, the actualist is also committed to the existence of something purely for explanatory purposes. The problem is that these are properties of which the actualist can give no account. They cannot be named either directly or even indirectly, but they must be there nonetheless (Lopston 1996, 430). Is this just an ad hoc solution? Intuitions differ on this point⁶, but it is clear that the actualist needs such entities to have a successful theoretical explanation.

⁶ Differing intuitions can often be equally strong. Peter Lopston calls this "a bizarre and deeply implausible thing to believe." (1996, 430)

Even so, it seems to me that the realist must concede that the principle of ontological simplicity is in the actualist's favor. Ontological simplicity is not the only quality of a good theory; in fact, it is probably not even the most desired quality of a good theory. This is reflected in the standard phrasing of Ockham's Razor: one shouldn't multiply entities needlessly. Sometimes, it may be best to sacrifice ontological simplicity for the sake of a greater good, but just what is that good? Ultimately, theories are explanatory tools, and all of the features of a theory must be judged for their explanatory value. It is clear that we should not postulate any entities that serve no explanatory purpose. It may be, however, that a sacrifice of ontological simplicity is justified if, in Quine's terms, it can "pay its way," that is, if it provides a great deal of explanatory power. A simple ontology is no advantage if it cannot explain the evidence.

Explanatory Power

Here lies the apparent strength of modal realism. Every actualist account takes some notion of possibility as primitive. For Adams, a possible world is a maximally consistent world-story; for Plantinga, it is a maximally obtainable state of affairs, for Stalnaker, it is an instantiable world property. Consistency, obtainability, and instantiability are all defined in terms of possibility. So, in all three cases, there must be some notion of possibility that cannot be explained. Other candidates that have been proposed for possible worlds suffer from the same defect. For instance, Nicholas Rescher, has proposed analyzing the possible in terms of the conceivable(Rescher 1979). Conceivability, however, is clearly just another modal notion So, actualist theories suffer from an explanatory shortcoming. A theory can never explain what it takes as primitive. Adams cannot explain consistency, Plantinga cannot explain obtainability, Stalnaker cannot explain instantiability, and so on.

Realism, on the other hand, makes no explicit appeal to any modal concepts. The modal realist explains possibility *not* in terms of things that *might* exist, but in terms of things that *do in fact* exist. A possible world is not a set of propositions that could all be true, but is a world that in fact exists. So, it seems, that all of the explanatory benefits that were discussed earlier can be gained only by paying the ontological price of modal realism. (Lewis 1986, ch. 3)

The price is too high, and the gain too small; for modal realism is a philosophical version of the emperor's new clothes. A closer examination reveals that realism must also assume some modal concept as primitive, although the realist is much better at hiding it.

There are several reasons that modal realism is committed to a fundamental modal concept. The first is that the other possible worlds, in order to ground the notion of possibility, must serve as representations, that is, other possible worlds must represent some feature of the actual world. Lewis has criticized representation in actualism as an inherently modal notion. (Lewis 1986, 151) Although realism does not explicitly portray itself as representational, it is as committed to representations of modalities as actualism is. The possible worlds of modal realism cannot in themselves serve as possibilities of the actual world. That things might have been different in the actual world is a feature of the actual world, not a feature of some other possible world. So, if another possible world is to serve as the ontological ground of possibility, then it can only do so because it represents some feature of the actual world. What could this feature be? The only candidate is a possibility (see Bigelow and Pargetter 1987; and especially Hymers 1991).

Susan Haack suggests that Lewis is committed to a primitive modal notion when determining which worlds are possible. She argues that the only available method for Lewis is to appeal to a notion of complete and consistent description. (Haack 1977, 425-429) The realist can simply respond that the possible worlds are the ones that exist. Ones that do not are impossible. Why think that this is the case? It cannot be because it is impossible for such a world to exist. That involves a modal presupposition. Similarly, how does one know that other worlds contain no impossible objects? The realist cannot deny that there are such things purely on the basis of their being impossible. (Shalkowski 1994, 677-678)

A related objection involves the *number* of possible worlds. Modal realism requires that there be a world for every distinct logical possibility. How can the realist be certain that there are enough worlds? Lewis' method is to adopt the principle of recombination. Remember that this assumes the thesis that there are no necessary connections between distinct existents. Any two spatiotemporally distinct entities can coexist. (Lewis 1986, 86-90) Denying a necessary connection, however, is tantamount to affirming a possible one. To say that any two spatiotemporally distinct tinct entities can coexist, is simply to say that it is possible for it to be the case.⁷

These points will also apply to the combinatorial theories proposed by Brian Skyrms (1976) and D. M. Armstrong (1986; 1989).

The realist could simply declare by fiat that there are enough worlds. In order to sufficiently ground modality, however, the claim that there are enough worlds is only relevant given the assumption that there could be no more. If there could be more worlds than there are, then there are possibilities that have no ground. (Shalkowski 1994, 679-680)

The following argument makes the point explicit:

- 1. Consider the proposition P: There are as many worlds as there are possibilities.
- 2. If P is false, then facts about possibilities are not grounded in facts about possible worlds.
- 3. If P is true, then either P is a necessary truth or it is a contingent truth.
- 4. If P is contingent, then it is possible that there are possibilities for which there are no corresponding worlds.
- 5. If it is possible that there are possibilities for which there are no corresponding worlds, then the realist must admit that there are in fact such possibilities.
- 6. If P is necessarily true, then its necessity cannot be grounded in the fact that there are as many worlds as there are. So, its necessity must be assumed.
- C. Therefore, the realist must concede either that there are more possibilities than worlds or that some notion of necessity is assumed.

The *counterpart* relation also involves a primitive notion of possibility. One might argue in the same fashion as above that the counterpart must serve as a representation of a *de re* possibility of the actual individual. The relation suffers from another problem, however. The counterpart relation is a similarity relation. If *a* is an individual in w_1 , then the counterpart of *a* in w_2 is the individual that resembles *a* more than any other thing in w_2 resembles *a*. If an individual has a counterpart in every possible world, then that individual exists necessarily. If an individual exists contingently, then there must be a world in which that individual has no counter-

part. Assume that a has no counterpart in w_3 . What reason is there for such a claim? It cannot be because there is nothing in w_3 that resembles a in any way. Any two objects will be similar to each other in some respect, even if it is just being a physical object, or existing in some world, etc. So, the only reason there could be for the claim that an individual has no counterpart in some world is that there is a degree of dissimilarity so great such that it is impossible for any two objects so dissimilar to be counterparts of each other.

For all of these reasons, modal realism is just as committed to some concept of modality that it cannot explain as actualism is. Some notion of modality must be accepted as primitive. Given the ontological cost associated with realism, then, there is no reason why one should adopt the position. Of course, ontological simplicity and explanatory power are not the only principles by which the theories must be evaluated. The principle of internal consistency remains to be applied. If theories of possible worlds fail in this respect, then any explanatory advantage at all is simply an illusion.

CHAPTER 4

PROBLEMS FOR POSSIBLE WORLDS

Introduction

All accounts of possible worlds assume some notion of possibility as primitive. At first, this seems more than enough reason to dismiss a possible worlds account of modality. As Lewis has stated, "it would gain us nothing to identify possible worlds with sets of sentences (or the like), since we would need the notion of possibility otherwise understood to specify correctly which sets of sentences were to be identified with worlds." (Lewis 1973, 86) This judgment is too hasty, as Lewis himself later admits, "Mind you, it would be no small advance if we could explain modality in general by taking one very special case of it…as primitive." (Lewis 1986, 155)

Circularity, in itself, does not mean that possible world theories are bankrupt as explanations. Any theory must have some primitive notions that cannot be explained without reference to them. So, if this were a sufficient reason to dismiss possible world theories of modality, then it would be a sufficient reason to dismiss all of theoretical science. The circularity charge would be devastating only if possible world accounts were taken to provide definitions of possibility and necessity in non-modal terms. Understanding is not limited to cases of explicit definition, however. Possible worlds accounts have quite successfully helped us to understand the precise logical relationships between modal claims. This success should be evidence that they have provided us with some understanding of possibility and necessity that we would not otherwise have (McMichael 1983, 53).

So, if it were simply a matter of taking some sense of possibility as primitive, then possible world theories would still be viable modal theories. Unfortunately, the problems are much deeper. Although successful on the surface. possible worlds theories cannot, in any way, satisfy the requirements for a theory of necessity and possibility. What are those requirements? Michael Dummett has aptly put them thus: "The philosophical problem of necessity is twofold: what is its source, and how do we recognize it?" (Dummett 1978b, 169) The problem of necessity has a *metaphysical* aspect and an *epistemological* aspect. The first concerns the truth of modal claims; the second concerns our knowledge of them.

A theory of modality must then accomplish two tasks. The first is to give some account of the truth conditions of statements concerning necessity and possibility. The second is to give some account of how we acquire knowledge of those truths. A theory might be considered successful even if it only partially accomplished these tasks. Unfortunately, possible worlds theories cannot accomplish either of these two tasks; therefore, they can provide absolutely no understanding of the nature of modality. The theoretical framework provided by a possible worlds account of modality is simply insufficient to ground modal truth.

In this chapter, I examine three problems. The first is the problem of modal knowledge. Here, I assume that possible worlds exist and that they serve as the ontological ground for modal facts. The problem of modal knowledge is explaining how, on these assumptions, we could ever have knowledge of a non-trivial modal fact. The second problem is the problem of relevance. This problem assumes that possible worlds exist, but questions their relevance to the truth of statements using modal operators. The third problem, the problem of coherence, is the most serious. As we examine this issue, we find that possible worlds can in no way illumine our understanding of modalities. This is so because the very notion of a possible world is fraught with irreconcilable paradox.

The Problem of Knowledge

There are certain statements containing modal operators that we very plausibly claim to know. Examples include 'If a statement is true, then it must not be false,' and 'It is possible to memorize the entire text of Barnes' *The Complete Works of Aristotle*.' One of the important issues facing the philosophical analysis of modalities is how we come to know such truths. How we come to know them will depend, to a great extent, on what makes them true. So, different modal theories will be committed to different accounts of modal knowledge.

There are some modal claims that present no problem at all for knowledge, or at least no special problem for knowledge. These are claims concerning the possibility that Bill Clinton is president and the possibility that the University of Oklahoma is in Norman. Since both of these are actually the case, then they are certainly possibly the case. These claims are not the type with which we are primarily concerned. Trivial modal facts are statements of fact that are constructed simply by prefixing 'It is possible that' to non-modal facts about the actual world. The interesting issue concerns the possibility of non-trivial modal knowledge (i.e., statements expressing claims of necessity or mere possibility).

For obvious reasons, problems of modal knowledge have traditionally been associated with Lewis' modal realism. For Lewis, what makes modal statements true are facts about worlds that are completely causally and spatiotemporally disconnected from us. Now, I claim to know many things about objects that are causally and spatiotemporally quite distant from me. For instance, I claim to know that Socrates died in 399 BCE. I grant that what makes that statement true, Socrates' actual dying in 399 BCE, is an event that is very distant. It happened on another continent, nearly 2,400 years before I was born. Even though the event is very distant, however, it seems that I can sketch a plausible, though not very detailed, account of how a causal chain might extend from that event to my belief that it occurred. In fact, if it were impossible to have such a causal chain, then the prospects of my knowing that proposition seem just as impossible. On the realist account, there can be no causal connection, no matter how distant, between another world and my believing some modal claim. Therefore, if modal realism is true, then I can have no modal knowledge; so if I do have such knowledge, then the truth conditions proposed by modal realism must be false (see Richards 1975).

The problem that is proposed here is certainly perplexing. On the realist account, there is a vast dichotomy between modal knowledge and ordinary factual knowledge. Consider the question of the existence of unicorns. In order to know if unicorns actually exist, it seems that it is necessary to experience a unicorn either directly or indirectly, e.g., by receiving an eyewitness report that was passed down from a reliable source through other reliable sources. Without such experience, the claim that unicorns exist is fantastic at best. So, the modal realist must deny that he knows that unicorns exist in the actual world, but he affirms that it is possible for them to exist in the actual world. Therefore, he must claim to know that they exist in some other possible world. For the realist, it seems to be much easier to have knowledge about some other possible world than about the one in which you exist.

These worlds themselves are existing things, made up of things like those that fill up our own world. How, then, can the realist claim to even know that these worlds are there? After considering this question, William Lycan points out that the only available response is that there must be such things if the theory is to be true. That is certainly not much justification. It might be some justification if the theory had a great deal more explanatory usefulness than its competitors, but, as we have seen in the last chapter, it does not (Lycan 1979, 295).

Lewis has considered these arguments and prepared a response. His immediate response is that this places modal knowledge in the same straits as mathematical knowledge. What are the objects of knowledge in the case of mathematical knowledge? A plausible case can be made that they are things that exist beyond the ken of our normal means of cognition. Should we then suspect mathematical knowledge? No, Lewis replies, "Our knowledge of mathematics is ever so much more secure than our knowledge of the epistemology that seeks to cast doubt on mathematics." (Lewis 1986, 109)

Some have pointed out in response that there is at least one significant difference between mathematical objects and the possible worlds of modal realism:

118

mathematical objects are abstract, while Lewis' worlds are concrete (Skyrms 1976). Maybe we can know mathematical truth *a priori*, since the objects of that knowledge are abstract. Lewis' worlds, however, are concrete, and all knowledge of the concrete is *a posteriori*. Of course, Lewis is not bothered by charges that rely upon a distinction between the abstract and the concrete, since he thinks the distinction rests on a dubious foundation (Lewis 1986, 81-86). I think it is fair to say, however, that, even though we may fail to make the distinction precise, it is a natural distinction to make. So, Lewis' tendency to dismiss such objections is not warranted.

Lewis must recognize this to some extent, for he argues that the relevant distinction between the *a priori* and the *a posteriori* is not one of abstractness and concreteness, but one of necessity and contingency. Assuming that modality follows the pattern of S-5, the modal status of a proposition is necessary. So, the contingent truths of the actual world can only be known *a posteriori*, but that those truths are possible is one that can be known *a priori*. (Lewis 1986, 111-112).

There are of course, several difficulties with Lewis' response. First, there are significant differences between mathematical claims and many other modal claims. The possibility of a human existing in a disembodied state is quite different from claims concerning number theory. Even if the latter can be known *a priori*, it is unlikely that the former could ever be. Second, Lewis ignores Kripke's (quite plausible) account of *a posteriori* necessary truths. Third, the response clearly begs the question. According to Lewis, one can know necessary truths because they are known *a priori*, and the proper account of the *a priori* truths is that which is neces-

sary. Lewis simply claims that the a *priori truths* and the necessary truths are coextensive, but provides no plausible argument for that claim.

Actualism, to a great extent, fares no better with respect to modal knowledge. On Stalnaker's account, the world-properties are part of the actual world, but they are no more causally connected to us than the worlds of modal realism. Adams and Plantinga seem to have an advantage, for worlds are either sets of propositions or able to be represented by sets of propositions. We do have methods for testing the consistency of sets (although we can only test for logical consistency, and, given Church's theorem, even then only to a point), so presumably these methods could in principle be applied to possible worlds. These methods, however, can only be applied in principle given the size of the sets involved. Each set contains, for every proposition, either that proposition or its negation. So, the only person in a position to test these sets for consistency is an omniscient being. We, being finite, could never actually test these infinite sets for consistency, and therefore could never have knowledge of modal truths.

Rescher's view of modality as conceivability suffers from a similar fate. Our failure to conceive of some state of affairs will surely not be sufficient evidence of its impossibility. What we conceive is conditioned by, and contingent upon, our historical location, cultural situation, education, etc. Also, our ability to conceive of some state of affairs should not be taken as evidence of its possibility, especially with respect to some of the more interesting modal claims. Consider Kripke's claim that its biological origin is essential to an individual. If so, it is impossible for the same individual to arise from two different parents, but it must be conceivable in some sense of the word. Surely, I can conceive of Bill Clinton having different parents than he did. The only reply, and it is a very plausible one, is that I really am not conceiving of the situation in which Bill Clinton had different parents, but only the situation in which different people are called his parents. But then, knowing that it is inconceivable will require knowing what the world must be like in order for that situation to obtain, and this requires knowing the full range of relevant facts. Again, only an omniscient being will be able to have knowledge of modal truth.

So, some versions of possible worlds theories are initially apparently more successful at explaining how we can have knowledge of modal truth; but all of them, in the end, fail to provide even a minimally plausible explanation. This shows, then, that even if possible worlds exist and serve as the ontological ground for modal facts, then we could never have knowledge of those facts

The Problem of Relevance

One reason why possible worlds are insufficient to provide modal knowledge could be because they are irrelevant to modal truth. The problem of relevance is the problem of demonstrating how, even if they exist, merely possible worlds could have anything to do with the modal facts in this world.

Again, this is most clearly a problem for the modal realist. The realist, as Peter van Inwagen points out, faces "the problem of explaining what these things would have to do with modality if there were any of them." (van Inwagen 1985, 119) If every world is causally and spatiotemporally disconnected from every other world, then nothing that happens in any world has any effect on any other world whatsoever. How, then, can something that happens in another world have any effect on what is possible in this world?

Michael Jubien suggests the following thought experiment. Imagine that God created everything that exists. If realism is true, then God created all of the possible worlds. Assume that it is true that Jubien could have been a taxi driver, then there is a world in which some person similar to (a "counterpart" of) Jubien is a cab driver. Now, imagine that God had, for some reason, decided not to create that world. Would it still be possible in that case for Jubien to be a taxi driver? It seems so, but then, the existence of that other world is irrelevant to the truth of modal claims in the actual world (Jubien 1988).

Lewis addresses this problem, but its not clear that he appreciates its force. He recounts the objection in this way: other worlds exist in the same sense as this world exists. To exist is to be actual, so other worlds actually exist. If the modal operators are to be understood as quantifiers, then they are quantifiers over things that actually exist. The realm of possibility involves not how things actually are, but how things might be different from how they actually are. Quantification over things that actually exist is only relevant to how things are, not how they might be.

Lewis, of course, rejects this argument, since he rejects the claim that other worlds actually exist. He sees the problem of relevance as arising directly from a rejection of the indexical analysis of actuality. The problem, then, is simply a terminological one, stemming from disagreement about the meaning of 'actual.' (Lewis 1986, 97-101)

122

The problem, however, only indirectly concerns the indexical analysis of actuality. It really concerns the realist conception of the nature of possible worlds and their relation to actual modal facts. For the realist, what *might* be the case in the actual world is determined by what *is* the case in another possible world. If something is possible in the actual world, then there is a possible world in which it is the case.

For the moment, let's grant that it is plausible to think that there are such worlds, and if something in fact occurs in another world, then it must at least be possible in this one. This, however, will not entail that the non-modal facts about other worlds determine the modal facts of this one. For example, consider the possibility of running a mile in less than four minutes. How do we know that such a thing is possible? The obvious reply is that we know it is possible because people have done it. So, what happens at a world is relevant to our knowledge of modal facts, although in this case it is a trivial modal fact. What happens in another world would then be relevant to our knowledge of the modal facts of this one, if we could have access to such worlds.

Now, imagine that it is May 5, 1954, and no one has yet run a mile in less than four minutes. Is it possible that someone do so? Of course it is, Roger Bannister will do it the very next day. It is not the case that it is possible because someone has done it, since no one has yet done so. So, intuitively, the possibility of an event occurring in this world does not require its having occurred in this world. Lewis replies, of course not, that it requires its having occurred in another world. If so, then, if it is possible in this world, it must be the case in another possible world. One can imagine that no one in this, the actual world, ever runs a mile in under four minutes; maybe those who had the innate ability were just too lazy to train hard enough. Then, the possibility of a sub four-minute mile requires that someone in another world run one.

But is this example really any different from the inter-world case? According to the realist, the other world is not any different from the actual world. If the possibility of an event does not require its occurring in the actual world, why must it require its occurring in another world? Why must it be the case, that for every way the world could be, there be a way that another world is? If it need not be the case, then it doesn't appear that the worlds of modal realism are even relevant to actual modality (see Jubien 1988, 304-305).

For some actualist conceptions of worlds, possible worlds are clearly relevant to modal truth, although the significance of the relevance is certainly far from clear. Consider, for example, Adams' account of possible worlds as maximally consistent sets of propositions. The most intuitive way of looking at these worlds is as maximal descriptions of a way the world might have been. If so, then there does seem to be a strong connection between possibility in the actual world and a possible world. What is not clear, however, is which direction this connection runs. Is this a real possibility for the actual world *because* there is such a possible world, or is there such a possible world *because* it is a real possibility? This is tantamount to asking if it is a fact because the description is true, or is the description true because it is a fact. My intuition sides with the latter. In any case, since the actualist explicitly analyzes possible worlds in terms of possibility, then possible worlds are relevant in the same sense as my having dinner tonight is relevant to my having dinner tonight. Relevancy is not always meaningful.

So, if possible worlds theories are to serve as enlightening theories of modality, then they must be relevant to the actual modal facts. The realist conception of possible worlds cannot guarantee the necessary relevance. The actualist conception can maintain relevance, but only in a trivial sense that cannot help us understand the nature of modality.

The Problem of Coherence

If the arguments in the previous section have any weakness, it is that they rely upon controversial intuitions. If one's intuitions differ, then one could rightly disregard these arguments. But the case against possible worlds theory cannot be so easily dismissed, for a case can be made that such theories are ultimately paradoxical.

In order to clarify what follows, it is important to understand the relation between possible worlds and sets of propositions. This relation is clearly apparent in the actualist theory of Adams, where possible worlds just are infinite sets of propositions. There is also a strong relation between worlds and propositions for Plantinga. In that account, a possibility is a state of affairs. For every state of affairs, there is a corresponding proposition. So, each maximally consistent state of affairs that is a possible world corresponds to exactly one maximally consistent set of propositions.

125

For Stalnaker and Lewis, the connection between possible worlds and sets of propositions is there, but it is not quite so obvious. Stalnaker and Lewis completely reverse Adams' approach and define propositions in terms of possible worlds. A proposition is the set of all worlds in which that proposition is true. So, even with this reduction of propositions to worlds, we can still construct a set of propositions for each possible world. Our world corresponds to the set of propositions that each have our world in their extensions. Any other possible world can be represented in this way.

So, any way that possible worlds are construed, there is an infinite set corresponding to each possible world. Either possible worlds just are maximally consistent sets of propositions, or they can be exactly represented as maximally consistent sets of propositions. Hence, if a possible world is actual, it contains all of the true propositions. There is no true proposition that is not a member of the set corresponding to that world. The concept of a possible world is coherent only if the concept of the set of all true propositions is coherent. As we shall see, no matter how innocent such a set seems at first, it quickly leads to paradox.

Several people this century have either stated or alluded to such paradoxes. The first is Bertrand Russell. In the closing paragraph of *The Principles of Mathematics*, he writes, "The totality of all logical objects, or of all propositions, involves, it would seem, a fundamental logical difficulty." (Russell 1996, 528) The second is David Kaplan, who presents a modified version of the argument given by Russell. The most thoroughgoing attack on the notion of a set of all truth has recently been published by Patrick Grim in *The Incomplete Universe*. Grim's primary goal is to show that the notion of an omniscient being is incoherent, but he states in several places that similar arguments could be developed showing that the concept of a possible world is equally incoherent. In the remainder of this chapter, I will present two arguments for the incoherence of a set of all truths. The first is a semantic argument based on the Paradox of the Liar. The second is a set-theoretic argument based on Cantor's Theorem.

The Paradox of the Liar

The Paradox of the Liar has the distinction of being the only logical paradox mentioned in the Christian Scriptures. In St. Paul's letter to Titus, he writes, apparently referring to Epimenides, that "It was one of them, their very own prophet, who said, 'Cretans are always liars, vicious brutes, lazy gluttons.' That testimony is true. For this reason rebuke them sharply, so that they may become sound in the faith...." (Titus 1:12-13a NRSV) Admittedly, as stated here, there is no paradox. To say that Cretans are always liars is ambiguous. It could mean either of the following:

A: Every Cretan sometimes utters a falsehood.

B: No Cretan ever utters a truth.

Neither is paradoxical, even if uttered by a Cretan. A cannot be false, for if so, every statement uttered by a Cretan must be true, including A itself. It can be true, so long as it is not one the falsehoods that the person occasionally tells. On the other hand, B cannot be true, since it was uttered by a Cretan. It can be false, since that implies only that some Cretans at least occasionally tell the truth, so long as it is not this particular Cretan on this particular instance. So, Paul has not stated a paradox, but the potential is clearly there. It fails because these statements do not necessarily guarantee self-reference. It is not difficult to state it in such a way that the paradox seems inescapable:

C: C is false.

C, if true, must be false, since it merely asserts its own falsehood. On the other hand, if it is false, then what it asserts is the case, so it must be true. So, C is both true and false, a paradox.

The paradox can easily be restated in terms of possible worlds. In all versions of possible worlds theories, each possible world either is, or corresponds to, some maximally consistent set of propositions. So, for any possible world w, there is a corresponding set, S, such that, if w were to obtain, then the members of Swould be true. Since S is a consistent set, all its members would be true; since it is a maximal set, there would be no truths that were not members of S. Now consider the following sentence:

D: D is not a member of S.

S, being maximal, must contain either D or its negation. If w were to obtain, then all of the members of S would be true, including either D or its negation. If Dis true, then it must be the case that D is not a member of S. Then, there is a truth that S does not contain, so S fails to be maximal. If D is false, then S must contain the negation of D. The negation of D is simply that D is a member of S. So if S contains the negation of D, it must also contain D. Then, S is inconsistent. So, S must fail at least one of the necessary conditions for representing a possible world. It either fails to be maximal, or it fails to be consistent. So, if there is no adequate solution to the Liar Paradox, then there can be no set that meets the conditions for a possible world.

The intuitive response is to reject sentences such as C as nonsensical, but the difficult task is explaining exactly why this is the case. C is certainly selfreferential, but many self-referential sentences seem to be perfectly meaningful, such as "This sentence contains five words." The paradox can also be generated with a pair of sentences, neither of which is explicitly self-referential: The next sentence is true. The previous sentence is false.

There have been many responses to the Paradox of the Liar. Each of these responses either fails to provide an adequate solution, or if successful, renders it impossible to form sets powerful enough to serve as possible worlds.

As Sainsbury (1995, 1) defines it, a paradox is an apparently untenable conclusion that follows from apparently true premises using apparently acceptable rules of inference. A solution to a paradox, then, requires that one show that the conclusion, one of the premises, or one of the inferential steps is really only apparently acceptable. Richard Kirkham (1992) outlines five criteria that an adequate solution must meet. First, it must be specific. It must show precisely where the false premise or faulty inference occurs. Second, it must contain no ad hoc postulations. It must provide an independent reason for rejecting a premise or rule. It cannot simply do so because it leads to a paradox. Third, it must avoid "overkill." The solution should not solve the paradox by ruling out sentences that are otherwise unobjectionable. Fourth, it should be complete, applying to all versions of the paradox. Finally, it should conserve our intuitions concerning truth and validity.

As Kirkham recognizes, it is likely that no solution will meet these criteria. It may be that any adequate solution must have some ad hoc element. It may also be that an adequate solution will conflict with our intuitions somewhat. Some solutions, however, will meet these criteria better than others. The goal is to minimize both the ad hoc nature of the response, and any conflict with intuitions (Kirkham 1992, 273-275).

The first response to the Paradox of the Liar is to deny that statements such as C have truth value. C only results in a paradox if one insists on the Principle of Bivalence, that every statement must be either true or false. The simple version of the Liar is avoided by postulating "gaps" in truth. If such statements lack truth value, then they fail to generate the paradox. A simple modification to C, however, shows that this response is too hasty:

 C^* : C^* is not true.

This is called the *Strengthened Liar*, and a theory which postulates truth gaps cannot cope with such sentences. The paradox follows, for if C^* is either false or has no truth value, then it is true. Denying that all statements have truth value will enable one to avoid Liar Paradoxes with respect to statements that ascribe falsity to themselves, but it will still succumb to statements that simply deny their own truth.

A more sophisticated response is to appeal to a distinction between sentences and propositions. The sentence is the linguistic entity and the proposition is the meaning which the sentence expresses. On this account, most, but not all, wellformed sentences express propositions. The truth value of the sentence is determined by the truth value of the proposition that it expresses; thus, a true sentence expresses a true proposition, and a false one expresses a false proposition.

The proposed solution to the Liar Paradox is to deny that such sentences express a proposition. If so, then both C and C^* are meaningless. Sentence C fails to generate a paradox, since it fails to express either a true or false proposition. At first, C^* appears to assert something true, if so, the sentence does produce the paradox. However, if C^* fails to express a proposition, then there is nothing whatsoever that it asserts. This is so also for sentences such as:

E: E expresses no true proposition

The tendency is to affirm the truth of E; but if E expresses no proposition, then E is meaningless and has no truth value.

The propositional response to the Liar Paradox allows one to avoid these paradoxes, but it is certainly not free from objections. There are those, following Quine, who strongly object to propositions in general. There are, I believe, independent reasons for including propositions in one's ontology. Even if there were not, if propositions truly solved semantic paradoxes, then the loss of parsimony would be a small price to pay.

The propositional response, although successful in blocking simple versions of the Liar, is an unsatisfying solution. With no independent reasons for believing that sentences like C and C^* fail to express a proposition, the propositional response is simply an ad hoc solution: C does not lead to paradox, because it fails to express

131

a proposition. It must fail to express a proposition, for otherwise, it would lead to paradox.

The ad hoc nature of the propositional response is also evident in these sentences:

D: D does not express a true proposition.

 D^* : D does not express a true proposition.

If D is either true or false, then the paradox follows. The propositional response avoids the paradox by claiming that D does not express a proposition. It appears, then that D expresses a truth, and the paradox follows. If D does not express a proposition, then it does not express anything, so it cannot generate a paradox. If so, however, then D^* does express a truth, but D and D^* are comprised of exactly the same words in exactly the same order. The advocate of the propositional response must affirm that D^* expresses a proposition while D does not, even though the two sentences differ only with respect to their indexical labeling. So, unless the advocate of the propositional response can provide a reasonable account of the difference between D and D^* , then there is no reason to think that the propositional response is a satisfactory solution to the Liar Paradox.

Another possible solution to the Paradox of the Liar is to adopt a hierarchy of propositions. This solution is first developed by Bertrand Russell while addressing a problem that has become to be known as "Russell's Paradox.

Russell's paradox is a paradox of classes. Intuitively, for every meaningful condition, there exists a corresponding class. For example, let H be the class of human beings and M be the class of published books written by Margaret Meade. All

132

and only human beings belong to the former class, and only Gone With the Wind belongs to the latter. Even the condition of being a round square has a corresponding class, the null class. One can also specify conditions for classes which themselves have other classes for members. The class of classes with more than ten members includes H as a member, and the class of unit classes includes M as a member.

Since classes can include classes as members, one might wonder if there are any classes that include themselves as members. It is obvious that most classes do not have themselves as members. The class of unit classes is not a member of itself, since it has more than one member. The class of classes that have more than ten members itself has more than ten members, so it must be a member of itself. So, it seems that one can specify the following meaningful condition, being a class that does not contain itself as a member. Therefore, there must be a class R: the class of all and only classes that are not members of themselves. Is R a member of itself? If it is, then it must meet the condition for membership, so it is not a member of itself. If it is not, then it does not meet the condition for R, so it must be a member of itself. If it is a member of R if and only if R is not a member of R, which is a contradiction.

Not every contradiction is a paradox. Most contradictions should simply lead one to deny the original assumption that led to the contradiction. An example is Russell's barber who shaves all and only those who do not shave themselves. The barber shaves himself if and only if he does not shave himself. The solution is to deny that there is such a barber. This solution is satisfactory only if we have no independent reason for believing there is such a barber, which, in this case, we do not. The original assumption that led to Russell's Paradox is the *Principle of Class Exis*tence (**CE**): that for every meaningful condition, there is a class whose members are all and only those things that meet that condition. Escaping the paradox requires one of two things. Either one must deny **CE**, or one must deny that the condition "being non-self-membered" is a meaningful condition (Sainsbury 1995, 107-111).

Do we have independent reasons for affirming CE? The principle does seem to be intuitively true. Even in the case of the barber, there is a corresponding class. In denying that there is such a barber, we affirm that there is a class that corresponds to this condition, the null class. So, the best solution is to maintain CE, but find a systematic way to restrict what counts as a meaningful condition.

Russell's solution was to appeal to a principle that he called the *Vicious Circle Principle*. He gave several different versions of this principle, together they express the idea that no totality can contain any members that are completely specifiable only in terms of itself (Sainsbury 1995, 124). The condition "being non-selfmembered" violates the Vicious Circle Principle, since it specifies classes only in terms of themselves. He held that any theory which does not meet the Vicious Circle Principle will succumb to the paradox.

In Appendix B of *The Principles of Mathematics*, he sketched out such a theory, the Theory of Types. This is further developed in the *Principia Mathematica* and in "Mathematical Logic as Based on the Theory of Types." (1908) In fact, most attempts to address Russell's Paradox have involved some version of his Theory of Types. The Simple Theory of Types sorts entities into various types at

134

different levels. The lowest type, type₀, consists of individuals. An individual is any entity that is not a class. The next type consists of classes of individuals, which is followed by classes of classes of individuals, and so on. Each successive type consists of classes whose members are objects of the type immediately preceding it. No class can contain members that are of the same type as the class. So, the Theory of Types rejects as meaningless any question concerning a class having itself as a member (Russell 1996, 497).

Strictly speaking, the Simple Theory of Types is not strong enough to address the Paradox of the Liar. In his later work, Russell developed the "Ramified" Theory of Types, which groups propositions into hierarchies. The first level contains propositions that refer only to individuals. The second level contains propositions that refer to first-level propositions, and so on. No proposition can contain as a proposition of the same or higher level as a constituent. Every meaningful proposition must specify its level in some way.

The Ramified Theory of Types, for some self-referential sentence, will either deny that the sentence expresses a proposition or deny that it is true. If the sentence fails to specify its level, then it is meaningless. If a self-referential sentence does specify its level, than it does have meaning, but it is just straightforwardly false. "This sentence is false" can be read as "This is a type_n sentence and it is false." Since it contains a type_n sentence as a constituent, it must be a type_{n-1} sentence. Therefore, it is simply false, and no paradox results.

The Ramified Theory of Types solves the problem that proved to be fatal to the propositional response. The sentence D will be rejected, since it contains itself

135
as a constituent. Type theory succeeds, however, where the propositional response fails. The Ramified Theory of Types explains why D and D^* are different, even though they appear to be exactly the same. D, being self-referential, must be either meaningless or false; D^* , however, is not self-referential, and is true. The two sentences, although linguistically identical, do not express the same proposition.

There are several mathematical reasons to reject the Ramified Theory of Types. For instance, identity cannot be defined, and certain bound theorems in mathematical analysis fail. One could, however, choose to accept these costs for the sake of a solution to the Paradox of the Liar. Before doing so, however, other hierarchical theories should first be considered.

One alternative is presented by Alfred Tarski in *Logic, Semantics, and Metamathematics* (1983). Tarski saw the Paradox as an insoluble feature of natural languages. Natural languages are semantically closed, that is, they contain semantic predicates, such as 'true' and 'false', that are predicated of sentences of that language. The Paradox of the Liar can only be generated in a semantically closed language. For this reason, natural languages are simply inconsistent. A semantically open language has no semantic predicates that apply to its own sentences, but it can have semantic predicates that apply to sentences of another language. The original language is the object language, a statement that predicates truth of a sentence in the object language is a statement in a meta-language. The meta-language is itself semantically open, so none of its semantic predicates will apply to its own sentences. A statement that predicates truth of a sentence in the meta-language must be a sentence in some other language, a meta-meta-language. On this conception of truth, there is no truth without qualification, there is only truth in some language. To say that a statement is true is really to say that the statement is true in L, where L is some object language. The phrase 'true in L' is not a component of the language L, so any sentence that uses the phrase 'true in L' is not a sentence of L, but of a meta-language. Consider the sentence "The German sentence 'Schnee ist weiß' is true." This is sentence of English, not of German. As it used here, the sentence 'Schnee ist weiß' is not even a German sentence, but is instead an English name for a particular German sentence. It just so happens that the sentence and its name are indistinguishable.

No statement, then, can truly predicate truth or falsity of itself. The language that the statement is in has no semantic predicates that can apply to its own sentences. The statement 'This sentence is false' is equivalent to 'This sentence is false in the object language.' This, however, is a sentence in the meta-language, and is therefore simply false. No paradox follows (Kirkham 1992, 278-282).

Tarski's account of truth has proved to be extremely powerful. But as a general solution to the Paradox of the Liar, it and Russell's Theory of Types both fail. Both theories provide sufficient solutions to the paradox, so long as one considers only simple versions that contain one self-referential claim. Remember, however, that the paradox can be stated using no self-referential sentences:

E: The following statement is true.

 E^* : The preceding statement is false.

On Russell's account, sentence E must be of a higher type than E^* , since it contains E^* as a component. Sentence E^* contains E as a component, so it must be

of a higher type than E. So, type_E > type_E > type_E. In order to analyze these sentences, the proponent of Tarski's semantic theory must claim that the language of E is the meta-language of E^* , and vice versa:

E: The following statement is true in L_{E^*} .

 E^* : The preceding statement is false in L_E .

Once this is done, however, the paradox ensues. If E is true in L_E , then E^* is true in L_{E^*} . Since E^* simply asserts that E is false in L_E , it follows that E is false in L_E , and E^* is false in L_{E^*} , and so on.

There are versions of the Liar's Paradox that cannot yet be solved by these hierarchical methods. Even if one of the two eventually succeeds, however, it will still not be enough to preserve the concept of a possible world. Adopting a hierarchy of truth may allow one to solve certain versions of the paradox, but in doing so, one gives up any possibility of a coherent set of all truths. For Russell, any statement concerning all truths must be of a higher order than any true statement, and therefore must itself be meaningless. For Tarski, any statement concerning all truths must be in a meta-language for which there is no higher meta-language. Such statements cannot be asserted as true. On both accounts, any concept of a set of all truths must be rejected as incoherent, and likewise, any notion of a possible world must be rejected.

A third hierarchical account of truth, proposed by Saul Kripke (1975), does allow one to coherently formulate claims about all truths. On Kripke's account, there are no hierarchies of propositional types, nor of truth-predicates, but of interpretations. Each level of interpretation assigns 'True' to more sentences than did the previous level. After a certain number of levels, every sentence that will receive a truth value will have one. The language of this level will also be able to predicate truth of its own sentences.

For example, consider the following overly simplistic case. Imagine a domain that includes the following things:

 $D = \{$ snow, 'Snow is white,' 'Snow is blue,' "Snow is white" is true,' "Snow is blue" is false,' 'This sentence is false' $\}$

The lowest level interpretation, I_0 , assigns a set of objects from the domain to predicates of the language. So for each predicate, there is a set that is the extension of that predicate. The interpretation also assigns objects from the domain to the antiextension of each predicate. In this case, I_0 makes the following assignments (ignoring anything that unnecessarily complicates the example):

 $I_0(\text{is white}) = \{\text{snow}\}$

 $I_0(\text{is not green}) = \{\text{snow}\}$

 $I_0(\text{is true}) = \emptyset$

 $I_0(\text{is false}) = \emptyset$

Notice that this interpretation assigns nothing to 'is true'. Each successive level will assign more to the semantic predicates than the previous one. Two rules govern such assignments. First, each interpretation assigns everything to a semantic predicate that the preceding level assigned. Every sentence that has a truth value on one interpretation has the same truth value on a successive interpretation. Second, for any sentence of the form 'x is F', interpretation I_n assigns that sentence to 'is true' just in case the previous level interpretation assigned the subject of that sentence to the extension of its predicate. If the previous level interpretation assigned

the subject of the sentence to the anti-extension of the predicate, then I_n assigns it to 'is false'.

So, in this case, I_1 makes the following assignments:

 $I_1(\text{is true}) = \{\text{`Snow is white'}\}$

 $I_1(is false) = \{ Snow is blue' \}$

Notice that there are some sentences in the domain that are not assigned any truth values on I_1 .

 I_2 makes more truth value assignments than I_1 . It makes these assignments: $I_2(is true) = \{$ 'Snow is white', "'Snow is white" is true', "'Snow is blue" is false' $\}$ $I_2(is false) = \{$ 'Snow is blue' $\}$

There is now only one sentence in the domain that has yet to be assigned a truth value. That is the Liar sentence, 'This sentence is false.' In this example, there are no truth values that will be assigned by any successive interpretations. Every sentence that can have a truth value now has one. So, the Liar sentence will never be assigned to either truth value. For it to be so assigned, there must be a previous interpretation I_{n-1} that assigns its subject to the extension of its predicate. Since its predicate is a semantic predicate, however, then there must be an interpretation I_{n-2} that assigns that subject to the extension of the predicate is also a semantic predicate, so the cycle continues. Therefore, at any interpretation, the Liar sentence has a truth value if and only if it is assigned to 'is true' at I_0 , but I_0 never assigns anything to 'is true'. So, on Kripke's analysis of truth, the Liar sentence never has a truth value. So, it is not the case that every sentence is either true or false. There are "gaps" in truth.

Kripke's analysis of truth avoids the problem plaguing the other two hierarchical solutions. At some point, there is an interpretation at which no successive interpretation makes any more assignments to truth values. Since there is a level at which all sentences that can have truth values do in fact have truth values, one can still speak meaningfully of all truths. The price that is paid, however, is the assertion that Liar statements have no truth value. Kripke's theory of truth denies bivalence, and as we have seen, must be susceptible to the Strengthened Liar. Instead of solving the paradox, we have simply returned to the attempted solution that was rejected first. (Kirkham 1992, 282-294).

The most extensive treatment of the Paradox of the Liar is Barwise and Etchemendy's 1987 book, *The Liar*. Building on two different conceptions of propositions, they provide two analyses of the paradox, a "Russellian" and an "Austinian" account. The Russellian account, although different from the Theory of Types discussed earlier, results in the same consequence. On the Russellian analysis one is forced to give up any notion of the set of all truths. Barwise and Etchemendy explicitly affirm this: "From this perspective, where our intuitive reasoning goes wrong is in thinking the world encompasses everything that is the case. Give up this assumption and the paradox is avoided: the Liar is not true, but this fact cannot be a fact in the world, a fact that can be truly described." (Barwise and Etchemendy 1987, 105) Later, they write, "What we give up on the Russellian view is the totality of world." (Barwise and Etchemendy 1987, 174) As with the Theory of Types and the Tarski Semantics, the Russellian solution effectively solves the Paradox of

the Liar, but does so by denying a totality of truth, which is required for a coherent notion of a possible world.

The Austinian account preserves the totality of the world, but, in the end, renders the notion unintelligible. On this analysis, propositions are about situations. The proposition expresses a state of affairs, and a situation is a set of states of affairs. A proposition is true just in case the state of affairs that it expresses is a component of the situation in question. The philosopher's favorite proposition, 'the cat is on the mat' expresses the state of affairs of the cat's being on the mat. 'The cat is on the mat' is formalized as $\{s; [O, c, m; 1]\}$, which says that situation s is of the type where the cat is on the mat. The proposition is true just in case the cat's being on the mat is a component of the situation, so, $\{s; [O, c, m; 1]\}$ is true if and only if $\langle O, c, m; 1 \rangle \in s$ (Barwise and Etchemendy 1987, 121-128).

Now, it is possible to formally treat an ambiguity found in the denial of English sentences. There is a subtle difference between denying a positive claim and asserting a negative one. Imagine that my very near-sighted aunt is visiting, and says, "Look, the cat is on the mat." I look in the direction she points and reply, "No." By my reply, I could have either denied that the situation is one of the type where the cat is on the mat, or I could have affirmed that the situation is one where the cat is not on the mat. In Barwise and Etchemendy's terms, the first is a denial of her statement, and the second is a negation. Formalized, the denial of $\{s; [O, c, m; 1]\}$ is $\sim \{s; [O, c, m; 1]\}$, which is true just in case $\langle O, c, m; 0 \rangle \in s$. The negation is

even if the cat is on the mat in another room. The negation however, requires that the cat not be on the mat.

The liar sentence is ambiguous in the same way. If it is a negation, then it is the assertive liar, $f = \{s; [Tr, f; 0]\}$. Otherwise, it is the denial liar, $d = -\{s; [Tr, f; 1]\}$. The first claims that s is a situation where the proposition is false; the second claims that it is not the case that s is a situation where the proposition is true. W is a model of the world that is both maximal and consistent. A proposition p is true just in case $\langle Tr, p; 1 \rangle \in W$, and false just in case $\langle Tr, p; 0 \rangle \in W$.

If the assertive liar is true, then the state of affairs that it expresses must be part of the situation that it is about, hence, $\langle Tr, f; 0 \rangle \in s$. Since every situation that obtains is part of the world, s is a subset of W, and $\langle Tr, f; 0 \rangle \in W$. This, however, is what it means for the assertive liar to be false. Thus, it cannot be true.

If the assertive liar is false, then $\langle Tr, f; 0 \rangle \in W$, and therefore, $\langle Tr, f; 0 \rangle \in S$, where s is some subset of W. As in the preceding paragraph, however, this is what it is for the assertive liar to be true, so it seems that the assertive liar can be neither true nor false. This conclusion only follows on the assumption that $\langle Tr, f; 0 \rangle$ is an element of the situation that the assertive liar is about; if it is false, however, it only follows that it is an element of some situation, not necessarily the same one to which it refers. So the assertive liar can be consistently considered false, so long as it is not part of the situation that it is about. A similar argument shows that the denial liar is true, so long as it is not part of the situation that it is about (Barwise and Etchemendy 1987, 164-170).

It follows, then, that both the assertive and denial liars cannot be about the maximal situation that we call the world. If they were, then they would be part of the situation that they were about. So, the conclusion that one should draw from this is "...while the world is as total as one could want, we cannot, in general, make statements about the world as a whole." (Barwise and Etchemendy 1987, 154) Any meaningful statement that appears to be about the entire world is really only about a proper part of the world.

In drawing this conclusion one renders much of possible worlds metaphysics unintelligible. The realist could no longer speak meaningfully of two worlds being spatiotemporally isolated. Spatio-temporal isolation must hold for all of the parts, not just some large set of the parts. One could not say that an impossible state of affairs is part of no possible world. For to do so, would only be to say that an impossible state of affairs is not part of some very large segment of a possible world.

So, if the Paradox of the Liar cannot be solved, then all possible worlds metaphysics fail. Even if the paradox can be solved, possible worlds metaphysics fail. No proposed solution can solve the paradox while salvaging the theoretical framework required for a possible worlds metaphysics. The solutions proposed by Russell and Tarski render any notion of a totality of truth incoherent. Kripke's solution fails to solve all versions of the paradox. The best current solution, that proposed by Barwise and Etchemendy, renders the theoretical claims unintelligible. Granted, no current solution manages to preserve the intuitions necessary for possible worlds, but that does not mean that no future solution will. There are other reasons, however, to believe that there can be no coherent notion of a possible world.

The Cantorian Argument

The Paradox of the Liar shows that there can be no set of propositions that is both maximal *and* consistent, since such a set must fail to have either the Liar sentence or its negation. Some may yet find the argument unconvincing, since there does seem to be something suspicious about the sentence 'This sentence is false,' although as we have seen, there are more intuitive (and seemingly innocent) versions of the Liar. There is at least one other argument showing that a maximally consistent set of propositions is impossible, and it does not depend on any questionable self-referential statements.

Assume that T is the set of all true propositions. T is clearly an infinite set; since there are true propositions about each of the real numbers, T must have a cardinality at least as great as, and in fact, greater than, \Re . The conclusion will follow regardless of the size of T, so, for the sake of simplicity, I will represent T as if its elements are denumerable:

 $T = \{t_1, t_2, t_3, ...\}$

The power set of T is the set of all of the subsets of T:

 $\{\emptyset, \{t_1\}, \{t_2\}, \{t_3\}, \{t_1, t_2\}, \{t_1, t_3\}, \{t_2, t_3\}...\}$

Corresponding to each of these subsets of T is at least one (in fact, infinitely many) true propositions. For instance, $t_1 \notin \emptyset$, $t_1 \in \{t_1\}$, $t_1 \notin \{t_2\}$, $t_1 \notin \{t_3\}$, $t_1 \in \{t_1, t_2\}$, and so on. By Cantor's theorem, the cardinality of the power set of T is greater than the cardinality of T, that is, there are more subsets of T than members of T. If the cardinality is the same, then there is a one-to-one onto function f mapping every member of T to a member of the power set.

Consider the set T^* which is made up of all and only the elements of T that are not themselves elements of the set that is assigned to them by f. T^* is clearly a member of the power set of T since it is composed only of members of T. T^* , however, cannot be in the range of f, since it differs from every member of that range by at least one member. That is, consider some arbitrary element of T, t_n . If t_n is in $f(t_n)$, then it cannot be in T^* , since T^* includes only those elements that are not members of the set assigned to them by f; so again $f(t_n) \neq T^*$. Likewise, if t_n is not a member of $f(t_n)$, then it must be a member of T^* . So, there is no function that maps the members of T to members of the power set of T. There are more members of the power set than members of the original set.

Therefore, since there is a truth corresponding to each subset of T, there are more truths than there are members of T. T, however, is the set of all truths. So, there are more truths than there are truths. The original assumption, that T is the set of all truths, must then be false. So there can be no set of all truths (Grim 1991, 91-92). This argument is as ontologically innocent as possible. The truths that are members of T could be propositions, states of affairs, or sets of worlds. The argument is committed only to sets, and every account of possible worlds is so committed. One could presumably restrict naïve set theory so as to forbid the construction of T; but in doing so, one denies the possibility of any metaphysical theory that is committed to T. There is one particular Cantorian argument that has been discussed in the literature for some time and generally attributed to David Kaplan. (see for example Davies 1981, 262; Jubien 1988, 307-308) Recently, however Kaplan has published his own (official, I suppose) version in Kaplan (1994).

The possible worlds theorist seems committed to the following two claims:

1) There are at least as many possible worlds as there are propositions.

2) There are at least as many propositions as there are sets of possible worlds.

The first follows from the intuition that there are at least as many distinct possibilities as there are propositions. For any proposition, presumably, it is possible that God entertain that proposition at some particular time. The second claim follows from the position that propositions either just are sets of possible worlds, as in Lewis and Stalnaker, or correspond to sets of possible worlds, as in all other accounts.

In order for both claims to be true, the set of possible worlds must have the same cardinality as the power set of the set of possible worlds. This is impossible (as can be shown by a Cantorian argument of the same form as the one just given about sets of truths). If there are k possible worlds, then there are 2^k sets of possible worlds. Notice that this argument does not rely upon any one conception of possible worlds. It assumes neither Modal Actualism nor Modal Realism; it assumes only that there is some kind of correspondence between propositions and sets of worlds.

David Lewis objects to the argument for 1). If propositions are sets of worlds, then there are propositions that could not be contents of thought. He takes

mental content to be determined by the functional state of the thinker, and there are fewer possible functional states than there are propositions (Lewis 1986, 104-108). The response has a certain ad hoc nature. Earlier, Lewis talked about mental content as sets of possible individuals. But since that compounds the cardinality problem, he later adopts another theory of content. The proposed functionalist theory of content may solve the paradox, but any theory that limits the contents of thought to the realm of the countable will serve just as well.

Several philosophers have granted Lewis' objection and proposed ways to circumvent it. Jubien constructs a one-to-one mapping of propositions to the power set of all propositions (Jubien 1988, 307-309). Grim devises a way to construct a possibility for every ordinal number, so that the number of possibilities are too large to form a set (Grim 1997, 149). Kaplan sees Lewis' objection as demonstrating a serious problem for a possible worlds semantics of modal logic. The logic should not constrain the metaphysics. If the semantics requires adopting one metaphysical view and rejecting another, then there must be something wrong with the semantics (Kaplan 1994).

So there are several ways to avoid Lewis' objection. These responses, however, are not really necessary. There is a great deal of intuitive force to 1), and the burden of proof is on the possible worlds theorist to show that the principle is false, not on the objector to show that it is true. Regardless, even if Lewis' objection stands, the earlier diagonalization argument does not refer to any mental content whatsoever. In the end, the Cantorian arguments are simple, yet forceful. There cannot be enough worlds to sufficiently ground modal facts.

Conclusion

Dummett's "philosophical problem of necessity" provides two simple adequacy conditions for any successful modal theory. A theory must first explain what makes modal claims true, and then, explain how we come to know them. These conditions are obviously related. If a theory cannot explain what makes modal claims true, then it can give no account of how we come to know them, for it can give no account of the thing that is claimed to be known. If a theory cannot give an account of how we come to know modal claims, then it cannot explain how we can know the theory to be true, for the theory itself must make some modal claims. So, a theory must meet both conditions. A theory that gives no plausible answer to one question ultimately fails, regardless of how well it answers the other. Some possible world theories might remotely meet one condition, but no possible worlds theory meets both.

Modal realism fails to meet the epistemological condition. If modal realism is correct, then modal facts are grounded in things to which we have absolutely no epistemic access. If we can have no access to that which makes a modal claim true, then how can we ever claim to know them? Modal realism entails modal agnosticism. Modal actualism fares better, but only slightly. Modal facts are grounded in facts about the consistency of infinite sets. We do have methods for determining the consistency of sets up to a point, but, in this case, only an omniscient being is in a position to apply those methods. We should again be agnostic with respect to modal claims. The other condition is a metaphysical one, explaining what makes modal claims true. To successfully accomplish this, a theory must be both coherent and relevant. Possible world theories fail the coherence requirement. Both semantic arguments and set-theoretic arguments show that there can be no set of all true propositions. Any theory that is committed to the existence of such a set is fraught with paradox. As has been shown, actualist theories are explicitly committed, and realist theories are at least implicitly committed to the existence of such a set.

Possible world theories also fail to meet the relevance requirement. Again, modal realism fails to show that facts about other worlds are in any way relevant to what is possible or necessary in this one. Actualism may be able to show that facts about maximally consistent sets might be relevant, but cannot explain exactly why they are relevant.

So all possible world theories fail miserably to meet the adequacy conditions for a successful modal theory. A successful theory must explain modality in terms of things that are coherent, accessible, and relevant. We want to know how things could be; the most relevant and accessible place to begin answering that question is with the way things are.

CHAPTER 5

÷

THE ESSENTIALIST SOLUTION

Introduction

Possible worlds cannot provide the foundation for the metaphysics of modality because there are no such things. Does that mean that the phrase should be struck from the philosophical lexicon? Absolutely not. The terminology of possible worlds has proved extremely useful for discussing modalities and understanding entailment relations between modal sentences. I will continue to use the terminology when it helps to clarify arguments. One must understand however, that by the phrase "there is a possible world in which…" I mean nothing more than "it is possible that…."

At this point, the metaphysics of modality appears to be a hopeless endeavor. We seem to be strongly committed to the existence of modal facts. It still seems as if it must be the case that triangles have interior angles equaling two right angles, that Bob Dole might have won the 1996 election, that Bill Clinton could not have been an earthworm, and that square circles are impossible. It seems obvious to most that these claims are true, yet we have no satisfactory account of their truth conditions.

There are at least three available options. The first is to deny that these claims are objectively true. Modal facts are then simply facts about human psychology. The possible is what is conceivable. If our cognitive capabilities were different, then we would be willing to assert quite different modal claims. We believe the preceding modal claims are true simply because we cannot believe otherwise. Of course, this itself is a modal claim. So, it appears that, even on this account, there are objective modal facts, but they are facts about what a human being can conceive or believe.

The second option is to affirm the objectivity of modal facts, but claim that they are brute features of the world and cannot be subjected to any further analysis. It is then senseless to ask what makes the fact that Bob Dole could have won the election true. It just *is* true, and that is all that can be said about it. The problem with this response is that there are ample reasons that one could give to explain why one asserts certain modal claims. Maybe some, even many, modal facts are brute features of the world, but the ones that we have considered don't seem to be things that can have no explanation whatsoever.

The third option is to affirm the objectivity of modal facts, and analyze them in terms of some other feature of the world. This is the option that I will explore in this chapter. As I see it, the failure of contemporary modal metaphysics lies in its dependence on contemporary modal logic. In this case, logic has not served as a tool for metaphysics, but rather has been a conceptual constraint. I propose an analysis of modality in terms of an early distinction in the history of metaphysics, the distinction between essential and accidental properties.

Essentialism is the position that some properties are essential to an object and others are not. The history of essentialism in philosophy has generally paralleled the history of modal logic. That is, essentialism was a prominent doctrine in the ancient and medieval periods, and was generally disregarded from the modern period until the second half of the 20th century. As with many philosophical topics, there really is no single position called "essentialism." In fact, a cursory examination of essentialist thought in Plato and Aristotle reveals two different kinds of essentialism. The differences between the two are subtle, but important.

Aristotelian Essentialism

Aristotle is the foremost ancient proponent of essentialism. In fact, essentialism is so often associated with Aristotle, that it is commonly called "Aristotelian essentialism."¹ One should be careful, however, not to assume that every position labeled "Aristotelian essentialism" is Aristotle's essentialism.

Although the focus of this section is Aristotle's essentialism, he was by no means the only philosophical essentialist in the ancient period. Aristotle's essentialism has more in common with Plato's essentialism than it does with many essentialist theories that are now labeled "Aristotelian."

The topic of Plato's dialogue, the *Cratylus*, is the "correctness" of a name. In other words, what is it about a name that makes it the correct name for an object? Two positions were proposed. The first, defended by Hermogenes, is that naming is simply a matter of convention. Therefore, any name that a linguistic community agrees upon is the correct name. The other position, defended by Cratylus, is that the correct name reveals the underlying nature of the thing named. In the end, Soc-

¹ Of course, this does not mean that the phrase is necessarily a compliment to Aristotle, in fact, it seems to often have a certain pejorative air when it is used. That is the impression that I get when reading Quine, for instance (see Quine 1976c).

rates concludes that, at best, a name simply reveals what the originator of the name thought was the true nature of the thing, but the person could have been mistaken. So, the true nature cannot be known simply by analyzing the language used.

The important concept that Plato introduced here is that a thing has an essence, which is its true nature, and it has this essence independently of how it is perceived: "...it is clear that things have some fixed being or essence of their own. They are not in relation to us and are not made to fluctuate by how they appear to us. They are by themselves, in relation to their own being or essence, which is theirs by nature." (*Cra.* 386e, trans. C.D.C. Reeve)

Apparently, the essence is the proper answer to the "What is F-ness?" questions that are proposed by Socrates in the early dialogues. For example, in the *Euthyphro*, Socrates chides his interlocutor, saying, "I'm afraid, Euthyphro, that when you were asked what piety is, you did not wish to make its nature clear to me, but you told me an affect or quality of it, that the pious has the quality of being loved by all the gods, but you have not yet told me what the pious is." (*Euthphr*. 11a, trans. G.M.A. Grube) Knowledge of the essence of the thing is genuine knowledge of the thing. Knowledge of other features is merely knowledge of incidental characteristics of the thing.

Notice that Socrates will not be satisfied with just any definition of the word 'piety'. Instead, he requires a definition of the thing itself, a real definition as opposed to a nominal definition. The essence of a thing is expressed in its real definition. This is the first type of essentialism, what Kit Fine has called the "definition model." (Fine 1994, 2)

Aristotle's essentialism was motivated by the problem of change, a problem that troubled many philosophers of the ancient world. The problem of change is the problem of explaining how a thing can change and yet retain its identity. How can something change and yet remain the same?

In his work, On Generation and Corruption, Aristotle distinguished various types of change: coming-to-be, passing-away, motion, growth, diminution, and alteration. In motion, growth, diminution, and alteration, the object that is changed persists through the change. Aristotle wrote, "...there is alteration when the substratum is perceptible and persists, but changes in its own properties, the properties in question being either contraries or intermediates." (319b11) Aristotle's examples are of a body that is ill and becomes healthy, or a piece of bronze that is spherical but made to be square. In both cases, the body and the bronze persist, even though some of their properties change. These kinds of changes are accidental changes.

In cases of coming-to-be and passing-away, however, the thing does not persist through the change: "But when nothing perceptible persists in its identity as a substratum, and the thing changes as a whole (when e.g. the seed as a whole is converted into blood, or water into air, or air as a whole into water), such an occurrence is a coming-to-be of one substance and a passing-away of the other...." (319b15-18) These changes are substantial changes.

Here, Aristotle is appealing to our common-sense intuitions about change. After a severe storm, a tree might have lost many leaves and several branches; nevertheless, It is still the same tree. If the tree were to burn in a forest fire, the pile of

ashes that results would not be the tree. If the object survives the change, then it is an accidental change. If it cannot survive the change, then it is a substantial change.

Corresponding to the two types of change are two types of qualities. An accidental quality is something "which may belong or not belong to some self-same thing." (*Top.*, 102b8) If whiteness is an accidental quality of Socrates, then it is possible that Socrates is white, but it is also possible that Socrates is not white. A change in an accidental quality is an accidental change. If Socrates were to spend too much time in the sun and become pink, then Socrates survives the change, albeit with some minor discomfort.

Given this definition of accidental quality, an intuitive definition of an essential quality is a quality of an object that is not accidental. On this conception, if Socrates is in fact human, and humanness is not an accidental quality, then it must be an essential quality. If Socrates were to lose the quality of being human, then Socrates would cease to exist. Accidental and essential qualities are then defined in modal terms, with respect to what an object can or cannot have. This is the second type of essentialism, modal essentialism.

Modal essentialism is the type that is most often attributed to Aristotle. For example, it is modal essentialism that is attributed to Aristotle in Quine's "Three Grades of Modal Involvement." This is seen in his claim that Aristotelian essentialism allows sentences such as $\exists x (\Box Fx \& (Gx \& \sim \Box Gx))$ (Quine 1976c, 176).

In a 1973 paper, Baruch A. Brody developed what he called "an unabashed traditional theory of Aristotelian essentialism." According to Brody, with reference to the passages in *On Generation and Corruption* cited above, an object has a prop-

erty essentially if and only if it would go out of existence if it lost that property. An object has a property accidentally if and only if it could lose the property without going out of existence (Brody 1973, 354).

Stated simply, modal essentialism is the position that objects have some of their properties necessarily. As Ruth Barcan Marcus put it, "Aristotelian essentialism takes it that, if anything is a man or a mammal, it is so necessarily." (Marcus 1971, 190) Of course, this formulation is inadequate for several reasons. First, some might see the scope of the modal operator as ambiguous. It could be taken to mean either that necessarily, there is something that is F, that is,

(1) $\Box \exists x F x$

or that there is something such that it is necessarily F, that is,

(2) $\exists \mathbf{x} \Box F \mathbf{x}$

The first requires only that it be impossible that there not be something that is F, or every possibility includes an F thing. The second requires that there be some object such that it is impossible for that particular object to not have the property F. Of these two, then, the second is the essentialist claim. So, a modal essentialist claim is a *de re*, not a *de dicto*, modality.

Modal essentialism can be precisely formulated in this way: P is an accidental property of x if and only if x in fact has P, and it is possible that x lose P and xstill exist. P is an essential property of x if and only x in fact has P, and it is not possible that x not have P and x still exist.

Formulating this requires an existence predicate (Ex: x exists). The statement 'there is something that is essentially F' is then symbolized in this way:

$$(3) \exists \mathbf{x} (F\mathbf{x} \& \sim \Diamond (\sim F\mathbf{x} \& E\mathbf{x}))$$

Which is equivalent to:

$$(4) \exists x (Fx \& \Box(Ex \supset Fx))$$

There is something that is accidentally F' is formulated simply by removing the negation that precedes the possibility operator in (3).

Although modal essentialism is the type most commonly attributed to Aristotle, and continues to be the type held by most current proponents of essentialism, it cannot be the same as Aristotle's essentialism. The mistake that Brody and others make is strongly linking essential properties to substantial change. It is natural to think that every property which, if lost, would result in a substantial change is an essential property. If so, then essential properties are identified with necessary properties, a necessary property being one that an object has in every possible world in which it exists. If Socrates has a necessary property P, then it is not possible for Socrates to not have P. If Socrates lost P, then the resulting thing would no longer be Socrates. Socrates would simply no longer exist. On this analysis, every necessary property is an essential property.

Aristotle, however, believed that there are necessary properties that are not essential. In the *Topics* (102a17-30), he wrote,

"A property is something which does not indicate the essence of a thing, but yet belongs to that thing alone, and is predicated convertibly of it. Thus it is a property of man to be capable of learning grammar; for if he is a man, then he is capable of learning grammar, and if he is capable of learning grammar, he is a man. For no one calls anything a property which may possibly belong to something else, e.g. sleep in the case of man, even though at a certain time it may happen to belong to him alone... for it does not necessarily follow that if something is asleep it is a man (trans. W.A. Pickard-Cambridge).²

The ability to learn grammar is both a necessary and a sufficient condition of being human. Since it is a necessary condition, if a human were to lose the ability to learn grammar, that thing *qua* human would no longer exist. The ability to learn grammar is not a property that a human person could fail to have. It does not, however, indicate the essence of a thing, and is therefore not an essential property.

Gareth Matthews presents an equally surprising example. Aristotle thought that having angles that equal two right angles is a necessary property of triangles. It eternally belongs to all and only triangles. It is not, however, part of the essence of triangles (Matthews 1990, 255-256).

At this point, Aristotle's essentialism faces two primary obstacles. They both concern the relationship between substantial change and essential properties. The first is determining when substantial change occurs. Given the relationship between the essence and change, we may be able to fix the essence of a thing if we could fix the point at which substantial change occurred. This, however, begs an important question: how much change is too much?

² Given this, it is misleading to talk of essential properties according to Aristotle at all. The term 'attribute' would be more fitting. Nevertheless, I will continue to speak of properties in the contemporary sense.

The second obstacle is explaining why some necessary properties fail to be essential properties. If there is a correlation between substantial change and the essence of an object, then it would be natural to think that all necessary properties were essential ones. Denying this demands an explanation, and an account of the difference between properties that comprise the essence of a thing and those that are necessary but non-essential.

Fortunately, a closer examination of the *Topics* can help solve these difficulties. There are some passages where Aristotle clarifies what counts as a substantial change. For instance, at 125b37-39, he writes "...for it is impossible for a thing still to remain the same if it is entirely transferred out of its species, just as the same animal could not at one time be, and at another not be, a man."

If a human were to become a dog, or an apple, then the human would no longer exist. A change in species, with some qualification, is a substantial change. A thing can acquire a new species, as occurs according to Aristotle's theory of embryology. An embryo begins as a living thing, but not an animal. It then becomes an animal, but not a human. At the last stage, it is a human animal. What the thing cannot do is cross parallel species. The human cannot become a plant, or a member of another animal species (Matthews 1990, 257). In this, Aristotle answers the question of how much change is too much. If a change moves the object from one species to another, then it is a substantial change. If the change does not result in a change of species, then it is apparently accidental.

In order to know what properties count as essential properties, one must know the relevant properties that pick out a species. This is found in the definition

of the species. This is consistent with Aristotle's claim that the definition signifies the essence of a thing. (*Top.* 101b39) As Alan Code puts it, a universal is an essential property of a thing if and only if both the name and the definition of the universal can truly apply to the thing (Code 1982, 103). So, Aristotle does not subscribe to modal essentialism. Instead, Aristotle's essentialism is closer to the definitional model found in Plato.

Although one can casually speak of the definition of a thing, it is important to note that individuals do not have definitions. For Aristotle, only the species has a definition. The essence of the species just is its definition. Since the name and definition both truly apply to any member of that species, then the definition of the species signifies the essence of the individual belonging to that species So, every individual has an essence, since every individual belongs to some species, but there are no "individual essences" that apply only to a single individual and not to other members of the same species (Code 1982, 112).

In a sense, Quine was correct in saying that Aristotelian essences are independent of the language used to refer to a thing (Quine 1976c, 175-176). Otherwise, we could describe an object as a sitting man. If the sitting man stands, then it seems that substantial change has occurred, for there is no longer a sitting man. It is essential to the sitting man that he be sitting. The sitting man is an accidental unity, or using Matthews' term a "kooky object." In *Metaphysics* E 2, Aristotle denies (or at least comes close to denying) that there really are such things, calling them "obviously akin to non-being." (1026b21). In any case, the definition model of essentialism helps Aristotle deal with such accidental unities. Even if there is such a thing as the sitting man, it is not part of his essence to be sitting. As a man, he belongs to the species human. Therefore, only the definition of that species will be the essence of a particular human. Presumably, the definition of the species human will include nothing about sitting, since sitting is an accidental property of individual humans. Either these accidental unities have no essences at all, or they have essences only in a trivial sense (Matthews 1990, 258-259).³

Recognizing that the essence is found in the definition of the species helps solve the second difficulty, which is Aristotle's belief that there are necessary properties that are not essential properties. 'Triangle' is not defined (neither by Aristotle nor by us) as 'a figure whose internal angles equal two right angles.' Instead it is defined as 'a three-sided, closed, plane figure'. If the triangle were to lose the property of having angles equal to two right angles, it would no longer be a triangle. This is not because the property is an essential one, but because it would no longer be a three-sided, closed plane figure. If a person were to become incapable of learning grammar, she would not be a person; not because the ability to learn grammar is essential, but because she would no longer be rational.

Aristotle does not present an entirely clear account of how one should go about discovering essential properties. He seems to think that it is a job for science. In *De Anima* (402b21-403a2), Aristotle writes that knowledge of the essence of a thing is gained by studying all of the properties of the things that belong to a par-

³ Accidental unities are an interesting puzzle in Aristotle's metaphysics. Gareth Matthews has contributed the most to discussions on the topic (see for example, Matthews 1982). In light of this. Code's formula that is stated above needs to be qualified in order to exclude essences of accidental unities.

ticular species. We have the right essence, when the definition of the species enables us to derive the other properties of the individuals that are members of that species. A definition that does not allow us to discover the other properties, or at least to give some account of what they must be like, is false. So, it is not the case that something may be defined any way that suits ones interests. Some definitions are genuine and others are not; a genuine definition of the species expresses the real essence of a thing belonging to that species (Irwin 1988, 63).

This discussion of essentialism in Aristotle is admittedly far too simplistic. Not only are there necessary but non-essential properties according to Aristotle, but there are at least two different types of them (see Irwin 1988, 63-64, for a discussion of the complications involved). In any case, the basic features of Aristotle's essentialism can now be stated. For Aristotle, to be an essential property of an object, the property must meet three conditions: 1) it is a necessary property, 2) it is part of the definition of the species of the object, and 3) losing the property results in a substantial (species) change in the object. Discovering the essence of a thing is a matter for science. What we regard as the essence should be taken from the best scientific theory available.

As Platonic and Aristotelian philosophy continued to thrive through the ancient and medieval periods, essentialism remained a standard component of metaphysical theories. Essentialism, for obvious reasons, began to decline with the rise of nominalism in the later medieval period. One can find essentialist thought in the writings of the Rationalists, but very little in the writings of the British Empiricists. One exception is John Locke, in his distinction between nominal essence and real essence. One might see Locke following the definitional tradition of Plato and Aristotle. In *An Essay Concerning Human Understanding*, he wrote, "Essence may be taken for the very being of any thing, whereby it is, what it is." Of course, for Locke, that may be all that can be meaningfully said concerning the essence of a thing, since real essences are "generally unknown."⁴ (Bk. III, Ch. III, §15) Later, in *A System of Logic*, Mill advocated modal essentialism, treating an essence as "that without which the thing could neither be, nor be conceived to be" (Bk. 1, Ch. vi, §2). Contemporary essentialism is in this Millian tradition.

Contemporary Modal Essentialism

The development of modal logic in the second half of this century revived philosophical discussion of essentialism, primarily because the development of the formal semantics for quantified modal logic enabled essentialist theses to be formulated precisely. In 1970, Saul Kripke delivered a series of three lectures at Princeton that were later published as *Naming and Necessity* (Kripke 1972; and, with minor revisions, Kripke 1980). Kripke's work triggered a newfound interest in essentialism, and since then, several people have made cases for the plausibility of particular essentialist theses.

Unlike Aristotle, most contemporary philosophers who advocate essentialism favor a form of modal essentialism, in which every necessary property is an essential property. Many arguments for these modal essentialist theses have really

⁴ Locke's discussion of the distinction between nominal essence and real essence is notoriously difficult. The consensus seems to be that for Locke, there is no real practical difference between the two (see Ayers 1981; Odegard 1974; Uzgalis 1988; Vienne 1993).

only been appeals to intuition. For example, one often hears questions such as "You don't really think that Bob Dole could have been a potato, do you?" These intuitions are naturally fairly strong, but intuitions are often wrong. So, intuitions alone do not provide sufficient reason to hold any essentialist theses; further positive argument is required. In this section, I will examine the plausibility of three contemporary essentialist theses.

The Necessity of Origin

One essentialist thesis that has been proposed by Kripke is the essentiality of origin. In a famous passage, he considered the possibility that the Queen of England had different parents from those that she actually had, and found it inconceivable:

How could a person originating from different parents, from a totally different sperm and egg, be *this very woman*? One can imagine, *given* the woman, that various things in her life could have changed: that she should have become a pauper; that her royal blood should have been unknown, and so on.... And so it's possible that even though she were born of these parents she never became queen.... But what is harder to imagine is her being born of different parents. It seems to me that anything coming from a different origin would not be this object (Kripke 1980, 113, emphasis in original).

He continues to say that the matter from which an artifact is made is essential to the artifact. A particular wooden table could not have been made from a different type of matter such as ice or stone. The resulting object would not be identical to the wooden table. The table could not even have been made from a different block of wood than the one from which it was actually made. So, it is not only the type of matter, but the specific instance of matter itself that is essential. As has already been pointed out, these are simply appeals to intuition. Consider first the case of the essential origin of an artifact. Intuition may be the only support one can give for this claim, I have yet to see a plausible argument for it. In a footnote, Kripke offers "something like a proof":

Let 'B' be a name (rigid designator) of a table, let 'A' name the piece of wood from which it actually came. Let 'C' name another piece of wood. Then suppose B were made from A, as in the actual world, but also another table D were simultaneously made from C. (We assume that there is no relation between A and C which makes the possibility of making a table from one dependent on the possibility of making a table from the other.) Now in this situation $B \neq D$; hence, even if D were made by itself, and no table were made from A, D would not be B (Kripke 1980, 114n).

Although initially plausible, this argument fails. One can construct a clear counterexample by substituting a non-essential property for the property of being made from a particular piece of wood. Let table B actually be on the north side of the room. It is possible that there be another table, table D, on the south side of the room. B and D are clearly not identical. So, if D were on the south side of the room, and there were no table on the north side, D would still not be B. This argument implies that spatiotemporal location is an essential property of an artifact. That conclusion is highly implausible, unless one believes that every property that an object has is essential to it. Kripke's "something like a proof", if successful, is not only a proof for the necessity of origin but the necessity of any property by which we can individuate two objects (Price 1982).

As is usually the case, it is more difficult to show exactly why the argument fails than to simply show that it fails. The reason Kripke's argument is so plausible is that Kripke treats 'B' and 'D' as names which rigidly designate their referents. Since they are rigid designators, they pick out the same object in every possibility. The second possibility that Kripke considers, the one in which B and D both exist, is enough to show that 'B' and 'D' refer to distinct tables. So, of course, the thing named by 'D' is not identical to the thing named by 'B', even when B does not exist. Using names as rigid designators begs the question.

So, in order to avoid begging the question, one should refer to a table as simply the table made from A or the table made from C. In the second possibility Kripke describes, these descriptions obviously pick out different tables. Do they also pick out different tables in all other possible cases? Only on the assumption that being made from the same piece of wood is a necessary condition for the identity of a table. Therefore, there is no way to strengthen Kripke's argument without simply begging the question.

So, one must return to appeals to intuition. It seems right that a table made from ice or stone would not be identical to one made of wood. That intuition is perfectly plausible. It is not so obvious that a wooden table could not have been made from a different block of wood. Consider a table that is carved from a particular block of wood. Surely it is possible that the block of wood from which this table was actually carved itself be cut from a slightly different part of the tree. Imagine, then, a new block of wood which is comprised of 95% of the same wood as the original block. Could the same table not be made from this block of wood? Here, the intuition is not so clear. Kripke admits that there is a problem with the vagueness of the notion. The thesis of the essential origin of artifacts, therefore, needs to be weakened at least to state that a particular artifact could not have been produced

from matter that differs very much (an intentionally vague condition) from the matter from which it was actually produced (Forbes 1986, 10).

In the case of the essential origin of an organism, however, the intuition is very strong, and there are good reasons for holding some version of the thesis. Most of these reasons involve pointing out the consequences of the contrary position.

Here is a version of an argument by Colin McGinn. Suppose that the essentiality of origin is false. Then Bob Dole could have developed from a different pair of gametes. If so, then another person could have developed from the gametes which actually produced Dole. In this case, however, it seems as if the second person would have the greater claim to being identical with the actual Bob Dole (McGinn 1999b, 61).

Here is an argument from Graeme Forbes, with some modification. Consider three possible situations, W_1 (which actually obtains), W_2 , and W_3 . Imagine also three sunflower seeds, each of which can produce a sunflower plant. In each situation, the following occurs:

 W_1 : plant₁ develops from seed₁.

 W_2 : plant₁ develops from seed₂.

W₃: plant₂ develops from seed₁ and plant₃ develops from seed₂.

Imagine that W_1 and W_2 are as similar as possible, and that the plants in W_2 and W_3 that develop from seed₂ are as similar as possible to each other. The intention is to imagine that plant₁ in W_1 , plant₁ in W_2 , and plant₃ in W_3 have no intrinsic or spatio-temporal differences between them. Denying that this is possible is tantamount to admitting that some properties are essential. So, they are qualitatively identical. The question is this: which plant in W_3 is numerically identical to plant₁?

There are four possibilities in this case:

(a) $Plant_1$ is identical to $plant_2$.

(b) Plant₁ is identical to plant₃.

(c) Plant₁ is identical to neither plant₂ nor plant₃.

(d) Plant₁ is identical to both plant₂ and plant₃.

All four of these possibilities have undesirable consequences. If (d) is the case, then both plants in W_3 are numerically identical to the same thing in W_1 . By the transitivity of identity, then, they must be identical to each other. So (d) allows for two distinct yet numerically identical organisms.

Since there are no qualitative differences between plant₁ and plant₃ then there is no reason to think that plant₁ and plant₃ are not identical. If plant₁ and plant₃ are not numerically identical, then the failure of identity must be grounded in something. What could it be? The only possibility is a non-qualitative essence, but the person who denies essentialism will not opt for that. So, possibilities (a) and (c) must be rejected. Possibility (b) is the only remaining option, Upon adding another situation, however, (b) suffers the same fate as (a) and (c):

W₄: Plant₁ develops from seed₁, but is qualitatively identical to plant₂ in W₃.

So, now, plant₁ must be identical to $plant_2$ in W_3 , but that contradicts the earlier rejection of (a). The only solution is to reject the assumption that makes this whole scenario possible. That is the assumption that makes (a) and (b) possible, that

numerically identical organisms can develop from different origins (see Forbes 1986, 8-9; and Forbes 1997, 518-519).

Of course, there have been objections to the necessity of origin for organisms. The intuition that Kripke appeals to is that Queen Elizabeth could not have been the biological daughter of the Trumans. Price claims that this, even if true, implies nothing about the actual origin of the Queen. Price states that "Elizabeth II necessarily did not come out of the Trumans' no more specifies a circumstance of the Queen's creation than 'Elizabeth II necessarily was not born inside a nuclear warhead' and 'Elizabeth II necessarily was not a member of the Rolling Stones at birth." (Price 1982, 36) Presumably there are a number of reasons that it is impossible to be born inside a nuclear warhead; lack of enough room being one. That fact, however, tells us nothing about the actual circumstances of the Queen's birth. Why, though, would one concede that it is impossible for the Queen to be the biological daughter of the Trumans? The only relevant fact is that the Trumans were not her actual parents. So, this does imply something about the circumstances of the Queens origin; she must have had the parents that she did, unless it is possible that the Queen have no biological parents at all.

It may be possible that a particular individual had no biological parents at all. For instance, consider a zygote formed by a particular sperm and egg. It is conceivable that God or some scientists build a zygote that is qualitatively and materially identical to the first. The individuals that result seem to be identical. This case does not show that there is nothing essential about the origin of an organism, however. It shows that maybe it is not the particular gametes from which the organism developed that are essential, but rather, perhaps, the matter and the configuration of the matter. How much of the original matter is required? Possibly a considerable amount is required, but not all. Now, essentiality of origin for organisms appears to have the same vagueness as it does for artifacts.

Very often, essentialist arguments for the necessity of origin are dismissed because they are taken to be arguments that coming from the same parents is sufficient to guarantee identity (for example, see McKay 1986; and Salmon 1981). This, however, is not the thesis that Kripke maintains. Kripke is only committed to the view that origin is necessary for identity; he is not committed to its being sufficient.

One must pay close attention to the difference between the strong and weak theses of the essentiality of origin. On the weak thesis, the origin of an individual is necessary for identity. On the strong thesis, it is both necessary and sufficient for identity. The weak thesis implies that it is impossible for some actual organism to develop from different cells than the ones from which it actually developed. Even if the resulting organism would greatly resemble the actual one, the two would not be identical. In the strong sense, it is impossible for any other organism to have developed from the same cells from which some particular organism has actually developed.

Weak necessity of origin, that it is a necessary property of each thing that it originate from the cells that it did, can be formulated in this way:

$$(4): \Box \forall x \forall y (Oxy \supset \Box(Ex \supset (Ey \& Oxy)))$$

where Oxy: x originates from y, in the sense of developing from the same cells. To see why this correctly represents the weak origin thesis, consider that in order to be
an essentialist claim it is clear that the *de re* necessity is required; so this formula gets the main thrust:

$$(5) \forall x \forall y (Oxy \supset \Box Oxy)$$

In possible worlds terminology, this would mean that if an actual individual originates from something, then it originates from that thing in every possible world. For this to be true, however, both an individual that originates from something and the thing from which it originates must exist. Hence, the qualification with an existence predicate:

(6)
$$\forall x \forall y (Oxy \supset \Box (Ex \supset (Ey \& Oxy)))$$

Here, if an actual individual originates from something, then it must originate from that thing in any world in which it exists. Of course, the essentialist will not think that this formula is true of only actual individuals, but of all possible individuals. So, the following must also be true:

$$(4) \Box \forall x \forall y (Oxy \supset \Box(Ex \supset (Ey \& Oxy)))$$

In order to clearly show the difference between the weak and strong versions of essentiality of origin, it should be pointed out that (4) is equivalent to:

$$(7) \Box \forall x \forall y (Oxy \supset \Box \forall z (z=x \supset (Ey \& Ozy)))$$

Paraphrased, this states that, if anything x that exists in any possible world originates from something y, then for any other thing z that exists in a possible world, if x is identical to z, then z originates from the same thing as x.

Strong necessity of origin would then be this:

(8): $\Box \forall x \forall y : (Oxy \supset \Box \forall z (z=x \equiv (Ey \& Ozy)))$

The essentiality of origin that Kripke describes is the weak sense, and there are good reasons for holding such a view. The strong version of essentiality of origin is not nearly so plausible as the weak version. It is entirely implausible in the case of artifacts. Surely a chair could have been made from the same block of wood from which the table was actually made. The chair and the table would not be identical, however.

If the strong version is true, then anything that develops from the same origin as an organism is identical to that organism. In the case of a particular human, anything that develops from the same egg and sperm cells from which he developed is identical to him. It is possible, however, for the fertilized egg to undergo fission and produce two distinct individuals. Then the original person would be numerically identical with two distinct persons, once again creating a problem for the transitivity of identity. David Lewis, John Perry, and Harold Noonan have argued that persons, not just fertilized eggs, can undergo fission. This is understandably controversial, but if true, there is even more reason to deny strong essentiality of origin (Lewis 1976; Noonan 1983; Perry 1972).

If it is possible for an individual to develop from no biological parents whatsoever (the extreme version of a test-tube baby), then the following principle of essentiality of origin is the most plausible:

(EO): It is an essential property of a thing that it originate from substantial amounts of the same matter in substantially the same configuration.

This principle can be formulated in the same way as (4), recognizing that Oxy is now a vague predicate.

There are reasons to think that the particular origin of a thing, at least in some sense, is an essential attribute of that thing. In the case of artifacts, these reasons lie primarily in intuitions that are plausible but not overwhelming. In the case of organisms, these intuitions are even stronger, but some might still find them implausible. Later, we will see that there are good reasons for these differences in intuitions.

The Necessity of Constitution

A second essentialist thesis, proposed by Kripke, Putnam, and others, is that a chemical substance's fundamental physical constitution is essential to it. And so, Locke was right to state that "The real internal, but generally in Substances, unknown Constitution of Things, whereon their discoverable Qualities depend, may be called their Essence." (*Essay*, Bk. III, Ch. III, §15) Current essentialists, however, pose examples for which there is a great deal of agreement concerning their actual constitution. The two most common examples are gold and water.

Notice that gold and water differ from other things, such as human beings, dogs, and oak trees. The latter are commonly called individual substances; meaning that they can be individuated. For instance, one can go to the local shelter and count the dogs, but it makes little sense to go to the lake to count the water. One does not count the gold at the jeweler's, instead one counts the gold rings and gold watches. Rings and watches can be individuated. Gold, in itself, cannot. Gold and water are not individual substances, instead, they are kinds of indefinite "stuff," the names of which are generally referred to as mass nouns. Some of the properties that we attribute to these substances are clearly accidental. Gold is used as a monetary standard and a material for jewelry and decoration. Water is also used for decorative purposes in fountains and ice sculptures. These properties, however, are not the only properties of the two substances. Gold has an atomic number of 79, meaning that there are 79 protons in the nucleus of an atom of gold. A molecule of water is a compound of two hydrogen atoms and one oxygen atoms. Are these also accidental properties? In other words, is it possible for gold to have an atomic number other than 79, or water to be something other than H₂O?

Putnam's belief that these properties, the atomic number or chemical composition of a substance, are essential developed from a consideration of the meaning of natural kind terms such as 'water.' In 'Meaning and Reference' and several other articles, Putnam criticized Fregean theories of meaning, according to which sense determines reference, favoring instead a theory of direct reference. He also finds that this theory of reference has certain implications for essentialism (see also Putnam 1975; Putnam 1990).

Putnam asked his readers to imagine a planet called "Twin Earth" which, except for some crucial differences specified by Putnam, is exactly like Earth. One difference is that the liquid called 'water' on Twin Earth is not H₂O, but XYZ ('XYZ' is an abbreviation for a long and complicated chemical formula). At normal pressures and temperatures, however, XYZ is indistinguishable from H₂O. When visitors from Earth go to Twin Earth, they will assume at first that 'water' has the same meaning on Twin Earth as it does on Earth. When the visitors from Earth discover that the thing called 'water' on Twin Earth is really XYZ, they do not report that water is XYZ on Twin Earth, but that the thing called 'water' on Twin Earth is not water, but is instead XYZ. Since water is H_2O on Earth, for the stuff called 'water' on Twin Earth to really be water, it would have to be H_2O (Putnam 1993, 151-153).

The important thing to note here is that Twin Earth is simply a possible situation. So, this argument has modal consequences that both Putnam and Kripke affirm. Since water is actually H_2O , for the stuff called 'water' in any possible situation to really be water, it must be H_2O . Given that water is actually H_2O , it is necessarily H_2O . It is not known *a priori* that water is H_2O , it is discovered (Putnam 1993, 159-161).

Before discovering the atomic number of gold, a scientist might well assert that gold could have a different atomic number. If Kripke and Putnam are correct, this assertion does not express a metaphysical possibility, but an epistemological one. If gold does have the atomic number 79, then it is impossible that it not have it; if it does not have that atomic number, then it is impossible that it does. So, to say that gold might not have the atomic number 79 means only that we do not know what its atomic number in fact is (Kripke 1980, 123-125).

The conclusions that Putnam and Kripke draw here provide the defender of modalities with a response to some of the criticisms made by Quine. Even if the analytic/synthetic distinction fails, the concept of necessity is still meaningful. Quine's objection followed from an assumption that all necessary statements were analytic. If it is the case that there are necessary truths that are not knowable *a pri*-

ori, then it must be the case that there are some necessary truths that are not analytic. Even if analyticity falls, necessity survives.

So, if Kripke and Putnam are correct, then it is essential to water that it be H_2O , and essential to gold that it have the atomic number of 79. For the modal essentialist, however, since every necessary property is an essential one, these properties do not comprise the entire essence of gold. It may also be an essential property of gold that it dissolve in aqua regia, if this is a consequence of its microstructure (see McGinn 1999a, 52-55).

The *a posteriori* nature of the essentialist claims made by Putnam and Kripke has recently been criticized by Joseph LaPorte. According to Putnam and Kripke, we discover that a substance has a certain microstructure by studying it. In order to discover that water is H₂O, two things must be the case. First, there must be general agreement about the phenomenological properties of water. Before, we discover that water is H₂O, we must agree about what stuff counts as water. Second, there cannot be more than one substance that has those phenomenological properties. Imagine that H₂O and XYZ are both found on earth, and prior to 1750 were both called water. When we discover their microstructure, do we have any reason to say that one is water and the other not? If not, then we have no reason to say that water is essentially H₂O.

LaPorte points out that this 'imaginary' case is actual. It is not the case that water meets these conditions; therefore we could not have discovered that water is H_2O . Deuterium is an isotope of hydrogen. Instead of a single proton in the nucleus, it has a proton and a neutron. It has the same atomic number, but a heavier mass.

Deuterium oxide (D_2O) is phenomenologically similar to water. It is a clear, odorless, tasteless liquid. Is it water? H_2O and D_2O are sometimes called 'light water' and 'heavy water' respectively. Encyclopedia articles about water will speak of each. This leads one to conclude that there are two types of water (LaPorte 1996, 120-121).

This conclusion is hasty. LaPorte's argument is no less an appeal to intuition than Putnam's. The different names are in themselves evidence that we believe them to be different substances. Their phenomenological properties differ, they have different freezing points, D_2O does not sustain life, etc. An equally plausible conclusion is that when D_2O was discovered, scientists concluded that this substance that appears to be water is slightly different from water.

Mellor (1977, 307-308) and Salmon (1981, 176-192), however, have shown persuasively that the necessity of constitution does not follow from the theory of direct reference without adding some essentialist assumptions. The scientist, after studying the composition of water, declares "Water is H_2O ." The "is" could either signify predication or identity. If the former, then the scientist is simply saying that water has the property of being H_2O , but what basis does one have then to claim that it is an essential property? Water could be something else on Twin Earth. If the "is" signifies identity, then one does have a reason to think that water is necessarily H_2O . Putnam, however, seems to be treating this as a necessary identity claim, but this begs the question. Samuel Clemens is Mark Twain. Given that the two are identical, they are necessarily identical. If 'a' and 'b' are rigid designators, and a=b, then necessarily, a is identical to b. If "water is H_2O " is a true identity statement with rigid designators flanking the identity predicate, it trivially follows that water is essentially H_2O (see also Forbes 1985, 195-196).

If the 'is' in 'Water is H_2O ' signifies predication, then the claim that water is essentially H_2O may follow, but only with an implicit essentialist assumption. The scientist discovers that water is H_2O on earth, and rejects her former belief that the substance called 'water' on Twin Earth is water. This, however, only follows if one assumes that the microstructure of the substance is essential to it. So, Putnam's argument is successful only if one assumes either that there is a necessary identity relation between water and H_2O (which is equivalent to assuming that 'water' and ' H_2O ' are rigid designators), or that the water is essentially H_2O . It does not prove that water is essentially H_2O .

Putnam's Twin Earth accounts do show, however, that the belief that water is essentially H_2O is quite plausible. At best, the essentialist may have an appeal to intuition coupled with the lack of any plausible non-essentialist counterexamples. The value of an intuition in philosophical argument is directly proportional to its strength, and this intuition seems extremely strong. In fact, the intuition grounding essentiality of constitution is stronger than that grounding essentiality of origin. So, in the absence of any arguments against the essentiality of constitution, one is free to make the assumptions necessary to derive the desired conclusion. In fact, this seems to be what much of science does as it develops reductive identity theories. The preceding objections do not show that water is contingently H_2O , but simply that the claim that water is H_2O cannot be derived from a theory of meaning. One can reject the theory of direct reference and continue to maintain the related essentialist doctrines, which some essentialists in fact do (for example, see Forbes 1986).

The claim that the microstructure of a substance is essential to it is very plausible. Unlike essentiality of origin, essentiality of constitution is plausible in both the strong and weak versions. It is plausible to think that all and only water molecules are H_2O . There have been few arguments purporting to show that it is not essential. Some of these have been addressed here, others will be addressed in the following chapter. At best, we will see that objections to this claim are objections to other theses proposed by essentialists. These other theses, although related to the original claim, are different. The claim that a substance's microstructure is essential is not dependent upon these other theses.

So, then, the following principle of essentiality of constitution best captures that proposed by Kripke and Putnam:

(EC): It is an essential property of a chemical element or compound that it have the same microstructure that it actually has.

(EC) can be formulated thus, where S is a second-order variable ranging over kinds of substances generally referred to by mass nouns, and M is a secondorder variable ranging over types of microstructures:

(10) \Box "S" x ($Sx \supset \exists M(Mx \& \exists y(Sy \supset \Box(Ey \supset My)))$

Informally, this states that any sample of a substance in any possible world has some particular micro-structure, and any sample of that substance in any possible world has the same microstructure in all worlds in which it exists.

The Necessity of Set Membership

The most plausible thesis proposed by contemporary essentialists is that a set has its members essentially. Intuitively a set is simply a collection of objects, and the same collection of objects could not be a collection of different objects. Of course, this can be shown with a bit more rigor.

There are two ways of defining a set. First, a set can be defined by enumerating its members. For instance, the set of U. S. Presidents after 1980 is defined as the set containing Ronald Reagan, George Bush, and Bill Clinton. If the set is sufficiently small, and all of the members are known, then defining a set in this manner presents no difficulties. As the number of members in a set increases, however, it becomes increasingly difficult to specify that set simply by listing its members. There are also times when we want to specify sets in cases where we do not know all of their individual members. So, sets are often defined in terms of some (perhaps very complex) concept, often called a condition, that picks out all and only members of that set. In this manner, one can easily define infinite sets such as the set containing all of the even integers.

It is natural to think that there is a set for every concept that can be expressed. As was shown in chapter four, however, this intuition, called the Principle of Comprehension, leads to set-theoretic paradoxes such as Russell's Paradox. Russell solved this paradox by introducing a hierarchy of sets, which effectively restricts the Principle.

There is only one identity condition for sets. This is the axiom of extensionality; for all sets A and B:

A=B if and only if $\forall x(x \in A \equiv x \in B)$

This states simply that two sets are identical just in case they have all the same members.

The axiom of extensionality naturally suggests that a set's members are essential to it, but it does not state this itself. This requires a modal version of the axiom. One can show that the members of a set are essential to it using an argument similar to one discussed earlier for the essentiality of origin. Let S_1 be the set containing a, b, and c; and S_2 be the set containing d, e, and f. If it is not essential that a set have the same members, then it is possible for both S_1 and S_2 to be the set containing g, h, and i. If, so, it is possible for two distinct sets to have exactly the same members, and be indistinguishable from each other.

So, the members of a set are essential to the set. Again, the argument is not demonstrative, but it is plausible in both the strong and weak versions. Having the same members is both a necessary and a sufficient condition for identity of sets. Therefore, the essence of a set is expressed by the following:

(ES): It is an essential property of a set that it have the same members that it actually has.

This is formulated thus:

 $(11) \square X''z (z \in X \supset \square z \in X)$

Informally, (11) states that any two sets in any possible worlds are identical if and only if it is necessarily the case that they have the same members (for a thorough account of the essences of sets, see Forbes 1985).

These are three of the most popular theses asserted by contemporary essentialists. As has been shown, they vary with respect to their plausibility. Some find all three be plausible, others may find only one or two to be so. One might take these differences in intuitions to be a mark against essentialism. A successful essentialist theory, however, can easily explain why this is the case.

Evaluating Modal Essentialism

Two competing accounts of essentialism were found in Aristotle. They are the modal and definitional models. On the modal conception of essential properties, all and only necessary properties are essential. So, to say that Socrates is essentially human is to say that he is necessarily human. This statement is ambiguous, however. In its simplest interpretation, Socrates is essentially human if and only if the sentence 'Socrates is human' is necessarily true. This unqualified assertion implies that Socrates is not only necessarily human, but necessarily existing. In order for that sentence to be true in all possible circumstances, the name 'Socrates' must refer in all possible circumstances.

So, modal essentialism is usually qualified in one of two ways. The first is to make Socrates' being human conditional upon his existing. Then, to say that Socrates is essentially human is to say that it is necessarily true that if Socrates exists, then he is human., or,

 $(12) \square (Es \supset Hs)$

As Plantinga points out, since objects have no properties if they do not exist, this is equivalent to saying that it is impossible that Socrates have the complement of the property 'human' (Plantinga 1974, 60).

Another way that the original formulation can be qualified is in terms of the identity of Socrates. Here, Socrates is essentially human just in case it is necessarily true that anything identical to Socrates is human. Thus,

 $(13) \Box \forall x (x=s \supset Hx)$

These two formulations are equivalent. The assumption that they are not equivalent entails that it is possible that Socrates be both human and not.

The modal conception of essential properties is the conception almost universally adopted by contemporary essentialists. Plantinga adopts (12) as his definition of an essential property (Plantinga 1974, 60). Kripke never explicitly defines essential properties, but speaks of properties throughout *Naming and Necessity* in terms of necessity and contingency. Brody adopts an unusual definition of essential property, saying that an object has a property essentially if it has the property, has it in all possible pasts in which it exists, and has it in all possible futures in which it exists from any time that it has the property. Although different from standard modal conceptions, given its reference to possible pasts and futures, it is nonetheless modal (Brody 1980, 122).

Given the results of the previous chapter, no modal conception of essential properties will be sufficient to accomplish the required task, which is to answer Dummett's problem of necessity. Ultimately, a modal conception of essentialism fails to solve the problem; it simply ignores it.

Modal essentialism conflicts with our ordinary understanding of "essence." A popular dictionary provides three definitions of 'essence' that are relevant to philosophical essentialism: 1. The intrinsic or indispensable properties that serve to characterize or identify something. 2. The most important ingredient; the crucial element. 3. The inherent, unchanging nature of a thing or class of things (*The American Heritage Dictionary of the English Language*, 3d. ed., s.v. "Essence.").

These definitions capture our ordinary use of 'essence.' We speak of the essence of a book, that which captures the fundamental key ideas of the work. We also speak of the essence of a person, as in "Howard Hughes was in essence a recluse." By this we mean to be capturing some key concept of the nature of the individual. The ordinary understanding of essence is that which you must know to understand the nature of a particular kind of thing.

Some of the essential properties proposed by the modal conception will be essential in this sense, but many will not. Modal essentialists recognize this, for it often leads them to create a distinction between "trivial essential properties" and "non-trivial essential properties."

For instance, since it is necessarily true that two and two equal four, it will be essential to Socrates that he live during a time in which two and two equal four. It will be essential to Socrates that he be a member of the set containing Socrates. It is necessary that if Socrates exists, then the set containing Socrates exists. On most essentialist accounts, it is essential that Socrates be distinct from Mt. Rushmore. Surely one does not need to know this in order to understand the nature of Socrates, for if one did, one would have to know what Mt. Rushmore is to understand the nature of Socrates. So, understanding the nature of a thing would require knowing everything that is distinct from that thing (see Fine 1994, 4-5).

According to modal essentialism, being a necessary property is both a necessary and a sufficient condition for being an essential property. Modal essentialists

are correct in assuming that it is a necessary condition, but the previous examples show that it is not a sufficient one. Some necessary properties are simply not essential properties.

This is not the only problem, however. Otherwise, the modal essentialist could just appeal to the distinction between the trivial and the non-trivial, claiming that non-trivial essential properties correspond to the ordinary understanding of essence. The task for modal metaphysics is to explain what makes modal claims true. Essentialism cannot do so if essential properties are defined in terms of necessary ones. What is needed is a theory of essentialism that makes no reference to modality, but that is sufficient to ground modal concepts.

Definition and Essence

Since modal essentialism is not sufficient to the task, our only hope of solving the problem of necessity is, I believe, with the definition model found in Aristotle's essentialist writings. The key to Aristotelian essentialism, as it is found in Aristotle, is the notion that a non-linguistic entity can be defined. Beginning with this, one can develop a theory of essentialism that solves the problem of necessity and explains our differing intuitions concerning necessary properties.

There are several different kinds of definitions. One is a stipulative definition. Here, the person defining the term determines the meaning of the term. When a person simply stipulates the meaning of a term, it is pointless to claim that the person got it wrong. Dictionaries, however, do not stipulate meanings; they report meanings that are already in use. Their definitions, lexical definitions, do not determine meaning; the meaning of a term determines its lexical definition. A dictionary could simply get the definition wrong. The idea of definition that is required for essentialism is closer to the lexical model. The essentialist must maintain that a thing can be defined in a way that is not subjective, dependent upon the desires or preferences of the person providing the definition. Instead, the definition is determined by the way the object is.

Essentiality, then, with its dependence on definition, is related to analyticity. The key principles of essentialism are best explained by comparing and contrasting essential truths to analytic truths. A common understanding of analyticity is that an analytic truth is one that is true in virtue of the meaning of its terms. Given the definition of 'bachelor' it must be true that a bachelor is unmarried. The definition of the word tells us not only what the word means, but also something about those things to which the word applies.

Essential truths are like analytic truths, in the sense that they are expressions of definitions. The definitions, however, are not definitions of words or concepts, but of the very things themselves. On this account, every natural kind has a definition that expresses the very nature of the members of the kind. The definition of the kind expresses the properties that a thing must have in order to be a member of that kind. Some plausible examples have already been discussed, such as "water is H_2O ." Given this, the essence of a thing is defined as the minimal set of properties expressing the fundamental nature of a kind that are jointly sufficient to determine that any object instantiating them be a member of that kind.

This distinction between truth in virtue of meaning and truth in virtue of nature is an intuitively plausible one. Imagine a child who asks why all bachelors are unmarried. The proper response is, "Because that is what 'bachelor' means." If asked why all samples of water are H_2O , one replies *not* that it is because that is what 'water' *means*, but that it is because that is what water *is*.

Corresponding to the distinction between analyticity and essentiality is a distinction between types of necessity. Analytic necessity is grounded in meaning, and metaphysical necessity is grounded in the nature of things. The difference can be seen in the scope of the modal operators. One would expect that analyticity, being truth in virtue of meaning, would result in necessity that pertains to sentences. Analytic truths are expressed as *de dicto* necessities. Essential truths, however, are not true in virtue of meaning, but true in virtue of the nature of the thing. So, the corresponding necessity should not be *de dicto*, but *de re*. Again, this corresponds to common intuitions. Even though all are willing to grant the necessity of the sentence 'Bachelors are unmarried,' few are willing to assert the necessity of any particular person's marital status.

Notice that meaning is logically prior to analytic necessity and not vice versa. It is not the fact that 'a bachelor is unmarried' is analytically necessary that determines the meaning of 'bachelor.' Rather, the meaning of the term determines which sentences containing that term are analytic. Likewise, it is not the metaphysical necessity of the fundamental constitution of water that determines its essence, but essence that determines necessity (There is a detailed discussion of the relationship of analyticity to essentiality in Fine 1994).

If the definition of a thing captures its fundamental nature, then it cannot be the case that the thing continues to exist even though the definition no longer applies to it. As Aristotle maintained, a thing cannot change its essence and survive. Essential change is substantial change. Given this, to say that a statement is true in virtue of the nature of the thing is equivalent to saying that a statement is true in virtue of the identity of the thing.

Notice that a theory of essentialism that is based on proper definitions explains the differing intuitions concerning essentiality of origin, constitution, and set membership. The claim that the members of a set were essential to the set was found to be highly plausible and there have been few plausible objections to it. One can find more objections to the claim that the fundamental constitution of an element is essential to it, but these are usually directed to some related thesis, such as the theory of direct reference, rather than against the essentiality of chemical constitution. The essentiality of origin, however, was found to be very suspect in some cases; and even in the plausible cases, intuitions differed radically. Notice that in each of these cases, the agreement concerning the essentialist claim is directly proportional to the agreement concerning the definition of the thing involved. There is widespread agreement that a set is simply a collection of particular objects. There is also widespread agreement that water just is H₂O. A plausible reason for the disagreement concerning the necessity of a person's origin, however, is that there is little agreement concerning the proper definition of a person. If there were more agreement concerning the definition, then presumably there would be less disagreement concerning the respective essentialist claim.

Essence and Modality

If essences are to provide the ontological ground for metaphysical necessity, then there must be some criteria for the essence of a thing that makes no appeal to necessity. In other words, one cannot claim that modality is determined by the definition of a thing, then in turn claim that the definition is comprised of all of the thing's necessary properties. It has been proposed that this criterion be truth in virtue of the nature, or identity, of a thing. Not all properties that are necessary will be essential in this sense, although all essential properties will in fact be necessary.

It is true in virtue of the nature of the unit set containing Socrates that it contain Socrates. So, the property of containing Socrates is both an essential and necessary property of the set containing Socrates. It is not true in virtue of the nature of Socrates, however, that he be a member of the unit set containing Socrates. The property of being a member of the unit set is a necessary property of Socrates. It is presumably necessarily the case that a set exist if its members exist, so there is no possible circumstance in which Socrates exists, but fails to belong to the set containing Socrates. Therefore, being a member of the unit set of Socrates is a necessary, but not an essential, property of Socrates (Fine 1994, 4-5).

Belonging to the unit set of Socrates cannot be an essential property because it tells us nothing of the nature of Socrates. No description sufficient to identify the unit set containing Socrates can fail to mention Socrates, but no mention of the unit set will transform an insufficient description of Socrates into a sufficient one. The essence of a thing is what it is to be a member of the kind to which that thing belongs. The essence of a set is to have the particular members that it does. So, the essence of the unit Socrates is that it contain Socrates. The essence of Socrates, however, is to belong to the kind to which he in fact belongs. Thus, the essence of Socrates contains nothing involving sets.

Another type of property that is necessary but not essential is existence. If existence is a property, then it is by all accounts a necessary property. Existence is a necessary property of the maple tree in my front yard. It is necessarily the case that, if the maple tree exists, then it exists. It is not essential, however. There is nothing about the specific nature of the maple tree that includes its existence. Knowing that the maple tree exists does not help one differentiate it from the other trees in my yard.

These two examples show two ways that a property can be necessary but not essential. First, in the case of Socrates and the unit set of Socrates, being a member of the unit set is a necessary property of Socrates, but his having that property follows from the essence of the set, not from the essence of Socrates. Second, in the case of existence, it is a necessary property not only of the members of some particular kind, but of the members of any kind whatsoever.

It is natural to think that having some properties entails having other properties. For, instance, a thing's being red entails that it be colored. Having the property of being a three-sided plane figure entails having the property of being a geometric figure. Any triangle has the former essentially and the latter necessarily. So, a property that is entailed by any property in the essence of a thing is a necessary property of the thing. Just as some properties are entailed by others, some properties are precluded by others. The property of being a three-sided plane figure precludes having interior angles that equal 360 degrees. So, this represents an impossibility for triangles, since it is precluded by their essence. Once one has an account of the necessities and the impossibilities, one can then determine the contingencies. Having a side of any particular length is neither entailed nor precluded by the essence, so this represents a genuine contingency for triangles (Jubien 1993, 111-115).

So, knowledge of the essence of a triangle allows one to determine what is necessary, possible, or impossible for triangles. It is plausible to think that this is the case for other kinds of things also. With knowledge of the essence of a thing together with other relevant knowledge (knowledge of the causal laws for instance) one could derive many of the other properties that the object must have. For instance, knowing that water is essentially H₂O, one can presumably determine many of its other properties (see Copi 1954).

Individual Essences

Although I have been referring to "the essence of a thing," these essences are essences of kinds, not essences of individuals. The kinds must also all be natural kinds, not artifacts. Many essentialists maintain that there are essences of individuals, which differ from individual to individual of the same kind. Any theory of essentialism must examine the case for such essences.

An individual essence is the property of being a particular individual object. Instantiating the individual essence is both necessary and sufficient for being that very individual. One could have an account of qualitative individual essences, in which the identity of an individual is determined by a set of qualitative properties held by the individual. The difficulty is in imagining a plausible set of qualitative properties that are sufficient to determine individual identity. Most essential properties that have been proposed (species membership, chemical constitution, etc.) are not sufficient since they are shared by all members of a particular species or all instances of a chemical element/compound. Some essential properties that pertain directly to individuals have been proposed, such as Kripke's idea about the necessity of origin. These properties, in their most plausible form, can at best only provide necessary conditions for identity.

For these reasons, those who have postulated individual essences (Plantinga, Adams, and Lycan, for example) have generally maintained that they are nonqualitative essences, or haecceities. A haecceity is an individual essence that is not analyzable in terms of the qualities of an individual. The haecceity of Socrates is simply the property of being Socrates. It is thought by those that postulate individual essences, that such things are needed in order to speak meaningfully of modal claims referring to individuals.

For instance, if it is the case that it is possible that Bob Dole won the 1996 election, then there is a possible world in which Dole won the election in 1996. This only expresses a genuine property if Dole in the actual world is identical to Dole in the possible world. Qualitative properties are not sufficient to ground identity across possible worlds, but that identity must be grounded in something, so there must be a property of being Bob Dole that both persons have (for a defense of haecceities compared to other accounts of individual essence, see Lycan 1994, 95-99).

There are several reasons not to admit haecceities into one's theory of essentialism, however. First, being non-qualitative, they are mysterious at best. There is always an ad hoc air about postulating something about which nothing can be said.

Second, they don't appear to be required to do the job for which they are intended. Consider a basic modal intuition, that Socrates could not have been a performance of Beethoven's Fifth Symphony. If so, then there must be some reason that Socrateity could not have been instantiated by some performance of Beethoven's Fifth. If this is the case, then it must have something to do with the qualitative properties of both Socrates and the performance. The haecceity does not seem to be needed.

Third, such things are usually postulated within the framework of a possible world metaphysics. Since there is something fundamentally wrong with that framework, the materials used to build it are suspect, and may simply not be needed.

Fourth, although there may be certain questions that can only be answered in terms of individual essences, there is no plausible account of how one goes about discovering such essences. Discovering an individual essence is not like discovering that water is H_2O . The epistemological problem concerning individual essences is insoluble (Enc 1986, 403).

Fifth, there is little that is gained with individual essences that cannot be gained with an essentialism of natural kinds. When can we know that something is possible for an individual? We can know easily enough when either the individual has accomplished the thing, or the thing has been (or could be) accomplished by another of the individual's kind. Bob Dole might have won, because other humans have won. In this case, an individual essence serves no explanatory purpose what-soever. I must concur with David Wiggins that "the logical impropriety and the use-lessness of 'haecceitas' fully match the syntactic impropriety of the word." (Wiggins 1980, 120).

A complete dismissal of individual essences may be too hasty. There are certain singular modal propositions that cannot be grounded in an essentialism of natural kinds. For instance, most would affirm that the sentence 'Socrates could not have been Plato' is true. This claim, if indeed true, appears to require individual essences. Of course, these essences still have no empirical value, and are postulated simply to ground intuitions about such sentences. I remain suspicious of haecceities, but am in no position yet to render a final judgment.

Essences of Artifacts

Artifacts present peculiar difficulties for the essentialist for several reasons. First, it is notoriously difficult to formulate a precise definition of the term. One distinction commonly appealed to is the one between artifact and natural thing. An artifact is commonly understood to be something that is fabricated, such as a chair, painting, or computer. Natural things are things like humans, rocks, samples of gold, etc. The problem with this distinction is that many of the things that are ordinarily considered to be natural could conceivably be synthesized in a laboratory, and worse, some things that are only synthesized in laboratories are surely natural kinds (e.g. some of the trans-uranium elements).

Second, it is sometimes difficult to distinguish artifacts from natural things, given that the same object can be described in both natural and artifactual terms. The most famous example is Aristotle's clay and the statue, which appear to be two things that temporarily occupy the same spatio-temporal position.

Third, there is no agreement concerning the proper identity conditions for artifacts. Consider the case of the Ship of Theseus presented by Hobbes in *De Corpore*. Imagine that Theseus, while sailing the ocean, notices that a plank on his ship needs replacing. He changes the plank, and (environmental ethics not being what they are today) tosses the old one overboard. Is the ship still the Ship of Theseus? Intuitively, changing one plank does not make a new ship. This conclusion, however, traps us into a sorites paradox. After several years, and many changed planks, Theseus suddenly realizes that every plank that makes up the ship has been changed. Now, is this ship still the same Ship of Theseus? If yes, then what is it about the ship that makes it the same? It can't be any of its parts, because they are all now different. If it isn't the same ship, then at what point did it become different? How much change is too much change?

Many people think that the resulting ship is still the same ship. Imagine now that Theseus has sailed to Central America. He finds himself at the small isthmus that is now Panama. Not having the advantage of a canal, nor wanting to sail around South America, he decides to take an easier route. He takes the ship apart plank by plank and moves the lumber to the other side of the isthmus. When he reaches the ocean, he reassembles the ship. Is this the same ship? Most people still answer yes.

Now the problem worsens. Theseus decides that his ship needs a complete overhaul. He pulls into the dock, and instructs the dock workers to change every plank. As they remove each plank, they insert a new one. Being frugal people, though, they take the old plank to the next dock and assemble a ship there. They use the same planks in the same order as the original ship. Which one is the Ship of Theseus?

My intuitions are that the ship that undergoes partial, plank-for-plank, change is the Ship of Theseus. These intuitions, however, are really nothing more than "mere intuitions." Nathan Salmon discusses a further complication. Imagine that someone discovered the Ship of Theseus this year, and, having heard of the classic problem, decided to concoct a scheme to own the vessel himself. This person officially donates the ship to a museum of antiquities, but transports it to the museum piece by piece. Of course, as he removes a piece, he replaces it immediately. So, by the reasoning above, the ship that remains in his workshop is the Ship of Theseus. This, however, is absurd (Salmon 1981, 221).

These examples show that the identity of an artifact is determined by our particular interests in that thing. Most artifacts are defined functionally. A ship must meet certain functional criteria in order to be a ship. In Salmon's case, these criteria are not important. Our interests in that case are historical. It is important that it be those very planks. The identity conditions of artifacts are relative to our interests, so there is nothing that is objectively true in virtue of the identity of an artifact.

No artifacts, then, have objective essences. Each artifact, however, will be composed of something. For instance, the wooden chair is still a piece of wood that has a particular fundamental physical constitution. The stuff of which it is made does have an essence. Qua chair, it has no objective essence; qua natural object it does. The definition of a thing that determines its essence is the definition that is used in the best theory of natural science that pertains to that thing.⁵

Knowledge of Essences

The relationship between the essence, the real definition of a thing, and its necessary properties answers the first part of the problem of necessity. The second part, the epistemological question, remains. None of this has any real explanatory value, unless an account can be given of how we can come to know the essences of things.

The first condition that an essential property must meet is that it be instantiated by every member of that kind. Obviously, if there is something of that kind that fails to have the property, then the property cannot be considered part of the essence. In that case, there would be an object of the kind for which the definition of the kind did not hold. Notice that this is a necessary condition, not a sufficient condition for an essential property. It is conceivable that every member of a kind

⁵ Notice that similar difficulties may apply to some natural things such as rivers and mountains. These are things that also seem to be identified primarily with respect to our interests, so they also might have purely subjective essences.

instantiate an accidental property. For example, there was a time when kangaroos were only found on Australia and the surrounding islands. That doesn't seem to be true in virtue of the nature, or identity, of kangaroos. Presumably taking a kangaroo to another continent does not result in a substantial change.

To discover the essence of a thing, it is not enough to simply identify all of the properties that are common to members of its kind. Although it seems likely that we would be able to distinguish the essential from the accidental in many cases, there are surely some cases in which this would be difficult. What is needed is some criterion that can be applied. A plausible candidate is explanatory power.

The essence has explanatory value that the other properties of an object lack. Consider the freezing point of water and its chemical constitution. The latter explains the former in a way that cannot be reciprocated. Explanations that appeal to essences will be more successful in the long run. Likewise, predictions made on the basis of accidental properties are less reliable than those made on the basis of essential properties.

Knowledge of essences, then, is no different from scientific knowledge in general. One hypothesizes that some set of properties constitutes the essence of a kind. This hypothesis is successful if it allows one to explain other features of that kind and successfully predict future observations. The more phenomena that can be explained and predicted, the stronger the hypothesis. The hypothesis is particularly strong if no adequate explanation can be given for that kind's having those particular properties. The essence is that which explains, but cannot be explained by other properties of the thing. Knowledge of essences is not *a priori*, but is empirical knowledge (Brody 1980, 135-153).

.

CHAPTER 6

OBJECTIONS AND REPLIES

Introduction

In the previous chapter, I argued that essentialism in the Aristotelian tradition provides the ontological ground for modality. In order to accomplish this task, I proposed an essentialism with three important features:

- (1) It is based on a notion of a real definition, rather than necessity. In this way, essences provide the foundation for necessity, not vice versa.
- (2) It is an essentialism of kinds, not of individuals. Each member of a kind has the same essence.
- (3) It is an essentialism grounded in empirical discovery. Most essential truths are discovered through scientific study, rather than by simple *a priori* reflection.

Essentialism has been, and continues to be, an important component in philosophical thought. Such thinking, however, is certainly not limited to philosophy. Many of the arguments one hears in everyday, ordinary discussions are really disagreements about essences. Assertions of the form "It can't be a ..., if it doesn't have...." are made regularly. Presumably, the first blank contains the name of a kind, and the second contains its proposed essence. This apparent widespread commitment to essentialism is taken by some to be evidence for its truth. Others, however, might side with Garth L. Hallett, and see widespread essentialism as something pathological, similar to "varied forms of cancer (carcinoma, sarcoma, leukemia, or lymphoma, in skin, lungs, brain, prostate, cervix, colon, breast, blood, bone, ...)." (Hallett 1991, 3) As is the case for any position, essentialism has been the target of some misguided objections. Most of these were directed to the meaningfulness of essentialist claims. After the development of the formal semantics for modal logic, however, such objections were seen to be groundless. If nothing else, modal logicians have enabled essentialists to precisely formulate their particular theses. Other objections, however, cannot be so easily dismissed. This chapter focuses on two objections that present a significant difficulty for the kind of essentialism that I have proposed. The first concerns the contingency of natural kinds in biology; and the second, the ultimate relativity of essences. I believe that the essentialist can adequately respond to both of these objections. The opportunity to respond to such objections has an added benefit, for the responses help to clarify both the details and the benefits of the definitional model of essentialism.

Natural Kinds

As I have defined them, all essential properties are necessary, but not all necessary properties are essential. So showing that a particular property is not a necessary property is not sufficient to show that essentialism is untenable. If essentialism is true, it must be the case that *some* properties are necessary to particular things; however, since if there are no necessary properties, then neither are there any essential ones.

The position that some properties are necessary to the objects that instantiate them has really never been as controversial as it has sometimes been claimed to be. Of course, the non-controversial cases involve necessary properties that are admittedly rather trivial. For instance, all things are necessarily self-identical. It is certainly true in virtue of the identity of Socrates, that Socrates be identical to Socrates. If existence is a property, then no object can lose the property of existence without ceasing to exist. All instances of change in the property of existence are substantial changes.

Such trivially necessary properties, however, are certainly not good candidates for the kind of essential properties that are required by my account. With the failure of possible world metaphysics, the task for essentialism is to provide a foundation for modality. Trivially necessary properties can never be sufficient to ground modality with any reasonable degree of explanatory value. The only properties that can perform this function are ones that definitively state the nature of a particular kind of thing, and trivially necessary properties are far too general to do this. A successful theory requires necessary properties that are not instantiated by every existing thing. It requires that there be natural kinds, and that these kinds be objective features of the world. The essential nature of physical kinds, such as water and gold, has already been discussed and defended. Even many of those who are suspicious of essentialism in general, admit that it is plausible with respect to chemical elements (Sober 1993, 146). Biological kinds including that of human being, however, are certainly more controversial.

The most commonly proposed candidate for a non-trivial, biological, essential property has been species membership. The idea has a great deal of intuitive appeal: Lassie could not have been a cat, and you and I could not become paramecia. Kafka's *Metamorphosis*, although a good story, describes an impossible state of affairs. Gregor Samsa could not have become a cockroach. The metamorphosis is not simply ceasing to be one kind of thing and becoming another, it is one thing's ceasing to be and another thing's coming to be. One must concede that, at this point, the evidence for the impossibility of such a scenario comes solely from intuition. Others may have different intuitions. James Baillie has argued that it is conceivable that a human become a fly and maintain identity (1990, 187-188). Even so, most would grant that the intuitions supporting essentialism are stronger and more widely held.

Although these essentialist intuitions are again very widespread, there is growing doubt among philosophers of biology that species membership is essential. If so, given that species membership is one of the most plausible examples of a non-trivial essential property, it appears that a robust theory of essentialism is doomed from the beginning.

Muhammad Ali Khalidi has argued that biological kinds are not essential since an organism belongs to several different kinds in different stages of its development. This requires taking the stages of an organism, such as being a pupa or a larva, to be natural kinds. These are certainly not the candidates for natural kinds that have been commonly proposed by essentialists (Khalidi 1993).

Some arguments purport to show that essentialism of natural kinds is inconsistent with evolutionary theory. Natural kinds, according to one argument, are immutable. If natural kinds were not considered to be immutable, then it would be senseless to speak of them in essentialist terms. Species evolve, however, and thus could not be natural kinds.

This argument fails to show a conflict between essentialism and evolution. The immutability of natural kinds requires only that the kind have an essence that is unchanging. The members of the kind can undergo change, although such change is substantial change. For example, one can change a sample of one element into a sample of another element in a particle accelerator. The essentialist who views the element that a thing is as essential to it must maintain, however, that the first sample does not persist through the change. The evolution of species requires only the possibility that an organism and its ancestors be members of different species. This, in itself, does not establish the immutability of species.

Another reason to think that essentialism conflicts with evolution is that evolution is thought to be a slow, gradual process. On this view, a species evolves into a new species by undergoing many small changes over a long period of time. Since these changes are each miniscule, there is no non-arbitrary point at which species change occurs. Essentialism, however, is thought to require sharp boundaries between natural kinds.

Although plausible at first, this argument fails for three reasons. First, the claim that the essentialist is committed to a sharp boundary between natural kinds is questionable. If the view of speciation asserted by this argument is correct, then species membership may be vague. This, however, is not necessarily inconsistent with the essentialist's claim that the species of an organism is essential to it. At best, it shows that there are some organisms that do not clearly belong to any particular species. For those transitional organisms, it is possible that nothing specific could be said about their essences.

The argument makes another questionable assumption; namely that since species change is the result of many small changes, there is no non-arbitrary point at which species change occurs. This begs the question against the essentialist by assuming that none of these small changes are any more important than the others. If the account of essentialism that has been proposed is correct, then some of these changes will have greater explanatory value and predictive power than others.

The final reason that this argument fails is that it assumes that speciation is imprecise. This, however, presupposes a theory of biological taxonomy that is no longer standard. Other views of speciation allow the point at which species change occurs to be precisely fixed. Unfortunately, however, these alternatives form the basis for another objection to essentialism.

Joseph LaPorte has recently argued that essentialism of biological kinds fundamentally conflicts with biological taxonomy. LaPorte's argument is this: There are several different taxonomic schema, or ways that biologists divide organisms into kinds. On all of these ways, it is a contingent fact that any organism belongs to a particular kind. The claim, then, that an organism could not possibly fail to belong to that kind is false (LaPorte 1997).

What constitutes a species? Surprisingly, there is little agreement among biologists on this matter. There are generally three different ways of differentiating species in contemporary biology. The first is the interbreeding approach commonly associated with Ernst Mayr (Mayr 1970; Mayr and Ashlock 1991). Here, a species is a group of interbreeding organisms that are reproductively isolated from other groups. One way that reproductive isolation could occur is that, for various reasons, organisms develop different mate recognition systems (Paterson 1985). The second is an ecological approach. Here, a species is a group of organisms that is characterized by a historical lineage and occupying a particular ecological niche (see Andersson 1990; and Van Valen 1976).

Both of these accounts are thought to present a difficulty for the essentialist. Consider LaPorte's example. A small population diverges from a large group of organisms. The smaller group finds a new ecological niche, and adapts to a new lifestyle. After a period of time, the two groups cease recognizing each other as mates, and become reproductively isolated. On both the inter-breeding and ecological approach to speciation, the two are now different species. That this occur, however, is surely not necessary: "Yet this could be a plainly contingent matter. Had the members of the little branch not taken on a new niche, or had there not been reproductive isolation between them and the members of A, they would belong to species A, given the species concepts in question." LaPorte concludes that "Both the interbreeding approach and the ecological approach seem to distinguish species on the basis of features that are only contingently possessed by organisms. Hence, species membership or non membership would not seem to be essential if either account is right." (LaPorte 1997, 101-102)

The third method of differentiating species is cladism, advocated primarily by Hennig (1979). Here, speciation is determined solely by genealogy. A species is a historical lineage of organisms between either two speciation events or between a speciation event and an extinction event. A speciation event is any event that causes one species to give rise to another species. Whether an organism belongs to a par-
ticular species is determined solely by its place on the genealogical tree. So, the development of species can be represented by a genealogical tree. The trunk is the original parent species. When a speciation event occurs, a branch is formed. Each branch results in a new species. Since a species is the line between two speciation events, the species represented by the trunk does not continue to exist. So, at each branch, one species ceases to exist, and two new ones are formed, even if the members of one of the branches bears no difference in properties to the members of the original.

LaPorte is correct in stating that Cladism is inconsistent with standard essentialism (LaPorte 1997, 103). Imagine a lineage of organisms that comprise species A. If a speciation event were to occur at some point, then a branch occurs that results in two new species, B and C. It is plausible to think that the members one of these species, say B, are phenetically identical to the members of A. Had the event not occurred, for whatever reason, then species A would have continued. It is plausible to think that, in this case, some of the resulting organisms in species A would be numerically identical to those in species B. Species membership, then, is not essential to an organism.

So, if cladism is correct, species membership is not essential. Cladism strikes most biologists, however, as extremely implausible. The thought that every formation of a new species is also the extinction of an old one is very counterintuitive. This has led most biologists and philosophers of biology to reject cladism, even if they also reject essentialism (Kitcher 1989; Sober 1993, 161). In fact, the essentialist should reject any historical descent account of the nature of species, even if it is not cladistic. If a species is determined by descent from an ancestor, then the essences of biological organisms are likely to be unknowable. It is very difficult, if not impossible, to determine the exact lineage of an organism. The fossil record is just too difficult to interpret. If the criterion for species is purely historical, the essences of organisms will differ radically from the essences of chemical elements. One reason to grant the plausibility of essentialism with respect to chemical elements is that those essential properties are all intrinsic properties. If a species is determined genealogically, then its essence consists of relational properties which can serve no real function in causal explanations. This conflicts with the process for determining essences described in the previous chapter. At best, historical accounts of species are not theories of what species are, but of how they come to be.

The essentialist, along with the great majority of biologists, must reject cladism as the correct approach to biological taxonomy. LaPorte's argument is not sufficient to show that the other two approaches, interbreeding and ecological, are inconsistent with essentialism. At best, his argument shows that if either of these approaches is correct, then the circumstances that lead to the development of a new species are contingent. The environmental change that led to the new group did not have to occur, but it is a mistake to think that this is inconsistent with the claim that an individual organism is essentially a member of its species. It means only that it is contingent that any particular species exist, not that it is contingent that any particu-

209

lar thing be a member of its species. The essentialist denies the latter contingency, but happily accepts the former.

Even so, these approaches should be rejected for other reasons. Interbreeding implies sexual reproduction, and many species either never reproduce sexually, or do not do so consistently. So, the interbreeding criterion is rejected by most botanists (McKelvey 1982, 50). Some members of a species cannot interbreed with members of the same species, because of damage to the individual resulting from disease, accident, or a congenital abnormality. Hybrid animals present another difficulty. Some distinct species can produce offspring. These cases show that the ability to interbreed is neither necessary nor sufficient for membership in the same species (Wilkerson 1995, 113-115). The ecological approach fares no better for the essentialist, since it seems possible that two population groups occupying distinct ecological niches share all of their important intrinsic properties.

Fortunately, the essentialist is committed to neither the interbreeding nor ecological view of species. Essentialism with regard to chemical elements is plausible because we can know the microstructure of those elements. Until recently, we had little knowledge of the microstructural properties of organisms. Recent work in mapping genomes may help us to develop a better account of species. Some biologists have recently suggested that species are a result of variations on a set of fundamental forms, an idea that is quite satisfying to the essentialist (Gould 1980).

Of course, genetic identity cannot be a necessary condition for membership in the same species, or the only species with more than one member will be composed of identical twins. Few individual organisms are genetically identical to any

210

other organism. Nonetheless, many of these genetic differences are causally inert. The task will be to determine the features of the genetic structure that are important for formulating scientific explanations and predictions. It may be the case that many of the currently accepted species are made up of more than one natural kind. One must concede that there is much work to be done, but there is little reason yet to think that biological kinds are not essential.

Here, I have examined an argument purporting to show that essentialism is inconsistent with the actual taxonomic practices of biologists. Three approaches to taxonomy were found to conflict with essentialism, but a quick survey of the literature will reveal that, regardless of their consequences for essentialism, there are significant problems with each of these approaches. No single one of them provides an adequate account of species. At first glance, this seems to support essentialism, but it actually suggests a much stronger argument against essentialism.

One might grant that each of these approaches to understanding biological species is seriously flawed. The claim that they should be rejected, however, follows only from the assumption that there is some objectively true account of biological species. It could be the case that each of these approaches are flawed because there is no best account to be found. Sometimes, biologists are interested in the historical development and lineage of organisms. If so, then cladism and evolutionary taxonomies best serve those interests. At other times, biologists are concerned with reproductive populations, then the interbreeding approach is best. Sometimes, biologists are interested in how organisms fit within an ecological system, then the ecological niche approach bests serves those interests. No approach is

any better or worse than any other, it depends only on what one's particular interests are. This is a version of the strongest objection to essentialism: that essences are relative to the interests and concerns of particular human beings.

Relativism and Realism

As has already been said, Quine was correct when he defined essentialism as "the doctrine that some of the attributes of a thing (quite independently of the language in which the thing is referred to, if at all) may be essential to the thing, and others accidental." Quine gave no argument against the coherence of such a position, but simply called it a "metaphysical jungle." It is, apparently, a jungle which he could not imagine anyone desiring to go (Quine 1976c, 175-176).

Quine's statement, however, focuses on a feature of essentialism that has been the target of a significant objection. If essentialism is to be a substantive metaphysical theory with consequences for our understanding of the nature of reality, the essence of a thing must be an objective feature of that thing. If essences are purely subjective, then they might be sufficient to ground subjective claims of necessity, but not necessity itself.

The objection is best understood in light of Aristotle's distinction between substantial and accidental change. In making this distinction, Aristotle appealed to some common-sense intuitions that most of us have concerning change. After a severe storm, a tree could very well lose many leaves and several branches. Nevertheless, the tree, albeit somewhat changed in appearance, is still the same tree. On the

212

other hand, if the tree were to burn in a forest fire, the pile of ashes that would result would no longer be the tree.

The essentialist, then, makes a metaphysical distinction of properties based on this intuition concerning change. If this is to succeed, then this distinction between kinds of change must be objectively real. If it is not an objective distinction, then essentialism fails from the beginning. Although most people intuitively think that some changes are accidental and others are substantial, it might be just a matter of social convention that this is so.

Thinkers as diverse as John Dewey and C. I. Lewis have argued that there are no objective essences. If so, there really is only one kind of change, and there is no objective reason to distinguish different changes. We do have reason to call some changes substantial and others accidental; but the reason is purely subjective, and relative to our interests. Although Aristotle denies that substantial change occurs when the sitting man stands, if one's interests were solely in sitting men, and not men in general, then one might have a different conception of what is essential.

With airplanes, our interests do not lie in the number of wheels that they have. We do not distinguish between planes that have a certain number of wheels on their landing gear and those that have fewer or more. If someone were to remove one wheel from a four-wheeled plane, we would not say that the plane is now a different thing. If someone were to remove a wheel from a three-wheeled, humanpowered, land vehicle (i.e., tricycle), we would say that it has become a different kind of thing. a bicycle. Why do we make this distinction? Presumably, it is because the number of wheels that the plane has on its landing gear does not significantly affect the operation of the plane. Removing a wheel from a tricycle has a substantial effect on its operation, just ask any child who has just removed the training wheels from her bike. One of our primary interests in the case of cycles is the number of wheels they have. If we had an interest in the number of wheels that planes have, then we would be inclined to regard the loss of a wheel as a substantial change (Copi 1954, 180).

It might be the case that our considering some changes as important is purely conventional, relative to our interests. In fact, this is the case for artifacts. This explains our conflicting intuitions with respect to the Ship of Theseus. If we are concerned with a ship as something that provides waterborne transportation, then identity is preserved so long as the ship can fulfill its function (provided, as in the example, it remains similar in structure to the original). If the ship were to undergo a change that resulted in a loss of the ability to float, then that change would be regarded as substantial. If we are concerned with a particular historical entity, say the *Mayflower*, then it is not so important that it perform a function, but that it be comprised of some particular bits of material. If this material were to change, then the change would be regarded as substantial. So, whether an artifact undergoes a substantial or accidental change is relative to our interests concerning that artifact. Such changes are simply a matter of convention.

It is difficult, if not impossible, to believe that every change that we consider to be substantial is simply a matter of convention. The living thing in my backyard with bark and leaves ceases to exist once it has been burned. Brody maintains that this cannot simply be a matter of convention, for trees ceased to exist upon burning long before there were people to have conventions about their identity. He asks, with obvious sarcastic delight, if our existence after death is simply a matter of convention, then "why haven't we insured our immortality a long time ago?" Brody's response to the claim that essences are interest-relative may be ultimately question-begging. One could claim that he is simply retroactively applying those conventions of identity. Even so, one has to admit that Brody has simply stated a very common fundamental intuition (Brody 1973, 353).

A related objection is that our language forces us to make distinctions between substantial and accidental change, and therefore between essential and accidental properties. We think that instances of change are substantial when we are forced to call the thing by a different word. An accidental change is one for which the same word picks out the object both before and after the change. My bookshelf has six shelves. If I removed one, there would be only five. The reason I don't see this as a substantial change is because I don't have special words for various kinds of bookshelves. Substantial and accidental change, then, are relative to the language used to describe the change.

This objection fails for the same reason as the first. Our language is a reflection of our interests. Copi points out the well-known and well-worn example of the different words for snow in Inuit. The Inuit have many words for snow because survival requires them to make these distinctions. If change is relative to language, then it will also be relative to our interests. Since all change is not relative to our interests, some change cannot be relative to language (Copi 1954, 181-182). The distinction between accidental and substantial change is one that we find quite natural, and use in everyday life. Brody's example of human death is a very good one. We treat the death of humans differently from other kinds of change because we recognize it as a going out of existence. We also use the distinction in more mundane matters. I can alter my house in many ways, and the insurance company will still cover it. If I change the color of the house or the appliances, the insurance company will still consider it insured. If I convert the house to lumber, however, the insurance company will no longer cover it. The new thing is not identical with the first (Brody 1973, 354).

The anti-essentialist should respond, rightly, that nothing has yet been done to show that these cases of supposed substantial change are not just reflections of our interests. At best, Brody has shown only that some interests are harder to give up than others. We all, being human, have an interest in human life and in having shelter to protect that life. We organize the world into different categories depending on our particular interests at the time. Whether a change is regarded as accidental or substantial depends on how one has organized those categories. It should be granted, then, that many of our judgments concerning substantial change are interest-relative, and hence, many of our judgments concerning essences are likewise interest-relative. One should not infer, however, that all judgments concerning essences are subjective, for some of our most important interests can only be met if there are objective essences.

The long term satisfaction of any of our interests requires that we understand the world to the greatest extent possible for us. To accomplish this, we are constantly formulating explanations of our experiences of the world, and using those explanations to make predictions about future experiences. Our interests, then, will only be satisfied if these explanations and predictions are successful. The more successful our predictions, the better our chances of using them to satisfy our long term interests. Our explanations and predictions will be successful in the long term only if our categories, our ways of dividing up the world, approximate the truth.

Assuming that metaphysical realism is true, then most of the features of reality are independent of our beliefs, desires, and interests concerning them. If this is so, then the laws governing these features of reality are also independent of our beliefs, desires, and interests. These independent features of reality, however, relate to our interests, since, being part of reality, we must interact with them. So, we have a deep interest in understanding objective reality and the laws that govern it. If this interest is to be met, then it must be possible to discover those laws and make reliable predictions about future experiences. But this is only possible if those categories to which the laws are committed reflect the basic organizational structure of the world. So, a commitment to metaphysical realism entails a commitment to objective categories. If there are objective categories, then the things that belong to those categories must do so in virtue of some features that they possess. These are the features that determine the natural kinds. These features, in turn, simply comprise the essence of the kind. Therefore, as Wilkerson maintains, a commitment to metaphysical realism entails a commitment to essentialism (Wilkerson 1995, 66).

There is, here, an admitted assumption of, and commitment to, metaphysical realism. Of course, realism has been challenged on many fronts, but I do not find most these challenges particularly troubling. Dummett may very well be right in saying that it is not clear that a convincing defense of realism can be made. He continues to say, however, that "it is unclear whether the anti-realist's position can be made coherent" (Dummett 1978a, 24). Given the choice between being less than fully convincing and being incoherent, I would gladly choose to be unconvincing. Given that I propose essentialism as a necessary commitment of science, the objections to realism that trouble me most are those that claim that such a view is incompatible with modern physics.

The special theory of relativity shows that certain properties of a thing are relative to a particular reference frame. For instance, whether or not two events are simultaneous can depend on the relative speed of the person observing them to those things. Imagine a person, Jane, who is a passenger on a train traveling very fast from east to west. Jane sees two simultaneous flashes of light on the train, one at the east end of the train and one at the west end. Jane sees these flashes exactly as she is even with a signpost next to the train tracks. Leaning on the signpost is another person, John, who also sees the lights. Since John was standing at the same position on the track that Jane occupied when she saw the lights, he also judged them to be simultaneous. Since Jane is at rest relative to the lights, in order for her to judge them to be simultaneous, they must flash at the same time. Since John is moving west to east relative to the lights, in order for John to judge them to simultaneous, the light at the west end of the train must have flashed before the light at

the east end. So, there is no fact of the matter regarding the simultaneity of the flashes independent of some reference frame of an observer. A similar case can be made for the length, mass, etc. of an object. Whether or not an object has a particular length is again relative to some reference frame. So, *some* features of the world that we commonly think are objective are actually subjective, relative to an observer.

This does not entail that all features of the world are subjective. In fact, the relativity of simultaneity and length only follow from a commitment to the absolute nature of certain features of the world. In the example above, the relativity of simultaneity follows only from a commitment to the absolute speed of light. The reason the light rays traveling from the source at the west end of the train to John's eyes had a greater distance to travel than the ones traveling to Jane's eyes. Since the rays arrived to their respective eyes at the same time, either the flash happened earlier for one than for the other, or one set of rays traveled faster than the other. Concluding that simultaneity is relative requires a commitment to the absolute speed of light.

In fact, the special theory of relativity begins with an assumption that the laws of physics are the same in all reference frames. There are features of reality that are objective, and not relative. The relativity of certain features of the world requires the non-relativity of others. There are certainly some properties of things that we have mistakenly believed be objective, but there must also be some that are in fact objective.

CONCLUSION

This dissertation has been an attempt to solve Dummett's two-fold philosophical problem of necessity: explaining the ontological ground of modality, and the source of our knowledge of modal facts. I have shown that modalities have played an important role in the history of philosophy, from Aristotle through Leibniz. The rediscovery of modalities in the 20th century has stimulated some of the most important developments in recent philosophy. One cannot undertake even a cursory survey of current themes in western philosophy without encountering explicitly modal concepts such as validity, counterfactuals, and supervenience. These advances, however, are only as well-grounded as the metaphysics that underlies them.

Unfortunately, the standard accounts of modal metaphysics can explain neither the truth of modal concepts nor how we could come to know them. Possible world theories of modality suffer from a single fatal flaw: there can be no possible worlds. The concept of a possible world is as coherent as the concept of a set of all truths, and such sets are ultimately paradoxical. Both actualist and realist accounts of possible worlds are committed to such sets, and therefore neither can provide a satisfactory account of the ontological foundation of modalities. If another account of the metaphysical foundation of modality cannot be given, then it appears that all of these philosophical advances were simply apparent.

Fortunately, the key to the problem can be found in the writings of Plato and, especially, Aristotle. I have advocated a return to the Aristotelian notion of es-

220

sence. Essences are not simply necessary properties, but are the properties that comprise the real definitions of things. The essence is what must be true in virtue of the nature of a thing. Every natural kind has an essence that is not simply a reflection of our desires and preferences, but is an objective feature of reality. These essences comprise the real natures of the members of the kind, and in turn determine what is necessary, impossible, or possible for a thing of that kind. This is the answer to the first half of the problem.

To know what is necessary or possible, then, requires that one know the essences of natural kinds. This requires studying nature. As Aristotle thought, only those things that can be studied scientifically have essences. Locke was right to maintain a distinction between real and nominal essences, but wrong to think that real essences are in principle unknowable. Knowledge of essences, then, is the real goal of science. It requires that we formulate hypotheses concerning kinds and their essences, and test them for their explanatory and predictive value.

The primary objections to essentialism have come from contemporary empiricists. These objections, I believe, rest upon a mistaken view of essentialism as non-empirical. We do not just begin with a concept about what is essential, and force experience to fit our preconceptions. Such a method is doomed to failure in the end. Instead, we begin with a hypothesis about what constitutes a kind, and study that kind to find the most likely candidates for the essential properties of that kind. We then test these hypotheses for their explanatory and predictive power.

In the end, the essentialist can grant Quine's claim that no statement is immune from revision. As the claims of science concerning kinds and their essences are tested and revised, we will know increasingly more about the nature of things. As we come to know more about how things are, we will come to know more about how they must be. Increased knowledge of such necessities, in turn, brings increased knowledge of possibilities. Essentialism, then, is not the stagnant doctrine that it is often portrayed to be, but the real window to the range of possibilities before us.

•

LITERATURE CITED

Adams, Robert Merrihew. 1974. "Theories of Actuality." Nous 8:211-231.

. 1981. "Actualism and Thisness." Synthese 49:3-41.

- Alanenen, Lilli, and Simo Knuuttila. 1988. "The Foundations of Modality and Conceivability in Descartes and his Predecessors." In Modern Modalities: Studies of the History of Modal Theories from Medieval Nominalism to Logical Positivism, edited by S. Knuuttila, 1-70. Dordrecht: Kluwer.
- Andersson, L. 1990. "The Driving Force: Species Concepts and Ecology." Taxon 39:375-382.
- Aquinas, Thomas. 1955. Summa Contra Gentiles. Translated by Pegis, Anton C. Notre Dame, IN: University of Notre Dame Press.
- Aristotle. 1984. The Complete Works of Aristotle: The Revised Oxford Translation. Edited by J. Barnes. Princeton: Princeton University Press.
- Armstrong, D. M. 1986. "The Nature of Possibility." Canadian Journal of Philosophy 16:575-594.

_____. 1989. A Combinatorial Theory of Possibility. New York: Cambridge University Press.

- Ayers, Michael R. 1981. "Locke Versus Aristotle on Natural Kinds." Journal of Philosophy 78:247-272.
- Baillie, James. 1990. "Identity, Survival, and Sortal Concepts." The Philosophical Quarterly 40 (2):183-194.
- Barwise, Jon, and John Etchemendy. 1987. The Liar: An Essay on Truth and Circularity. Oxford: Oxford University Press.
- Beedle, Andrew. 1996. "Modal Fatalism." The Philosophical Quarterly 46 (185):488-495.
- Bigelow, John, and Robert Pargetter. 1987. "Beyond the Blank Stare." Theoria 53:97-114.

- Bochenski, I. M. 1940. "Sancti Thomae Aquinatis de Modalibus Opusculum et Doctrina." Angelicum xvii:120-218.
- Boland, Vivian. 1996. Ideas in God According to Saint Thomas Aquinas: Sources and Synthesis. Leiden: E. J. Brill.
- Broadie, Alexander. 1993. Introduction to Medieval Logic. 2d ed. Oxford: Clarendon Press.
- Brody, Baruch A. 1973. "Why Settle For Anything Less Than Good Old-Fashioned Aristotelian Essentialism." Nous 7 (4):351-364.

. 1980. Identity and Essence. Princeton: Princeton University Press.

Chisholm, Roderick M. 1963. "The Logic of Knowing." Journal of Philosophy 60:773-795.

. 1970. "Events and Propositions." Nous 4:15-24.

. 1979. "Identity through Possible Worlds." In *The Possible and the Actual: Readings in the Metaphysics of Modality*, edited by M. J. Loux, 80-87. Ithaca, NY: Cornell University Press.

- Code, Alan. 1982. "On the Origins of Some Aristotelian Theses About Predication." In *How Things Are*, edited by J. Bogen and G. E. McGuire, 101-131. Dordrecht: D. Reidel.
- Copi, Irving M. 1954. "Essence and Accident." The Journal of Philosophy 51:706-719.
- Curley, E. M. 1984. "Descartes on the Creation of the Eternal Truths." The Philosophical Review 93:569-597.
- Davies, Martin. 1981. Meaning, Quantification, Necessity. Boston: Routledge & Kegan Paul.
- Denyer, Nicholas. 1981. "Time and Modality in Diodorus Cronus." Theoria 47:31-53.
- Dummett, Michael. 1978a. "Truth." In Truth and Other Enigmas, 1-24. London: Duckworth.

_____. 1978b. "Wittgenstein's Philosophy of Mathematics." In Truth and Other Enigmas, 166-185. London: Duckworth.

- Enc, Berent. 1986. "Essentialism without Individual Essences: Causation, Kinds, Supervenience, and Restricted Identities." In *Studies in Essentialism*, edited by P. A. French, T. E. Uehling, Jr. and H. K. Wettstein, 403-426. Minneapolis: University of Minnesota Press.
- Fine, Kit. 1994. "Essence and Modality." In *Philosophical Perspectives*, 8: Logic and Language, edited by J. Tomberlin, 1-16. Atascadero: Ridgeview.
- Flew, Antony. 1986. David Hume: Philosopher of Moral Science. Oxford: Basil Blackwell.

Forbes, Graeme. 1985. The Metaphysics of Modality. Oxford: Clarendon Press.

. 1986. "In Defense of Absolute Essentialism." In *Studies in Essentialism*, edited by P. A. French, T. E. Uehling, Jr. and H. K. Wettstein, 3-31. Minneapolis: University of Minnesota Press.

_____. 1997. "Essentialism." In A Companion to the Philosophy of Language, edited by B. Hale and C. Wright, 515-533. Oxford: Basil Blackwell.

- Frege, Gottlob. 1980. *The Foundations of Arithemetic*. Translated by Austin, J. L. 2d ed. Evanston, IL: Northwestern University Press.
- Gould, Stephen Jay. 1980. "Is a New and General Theory of Evolution Emerging?" *Paleobiology* 6:119-130.
- Grim, Patrick. 1991. The Incomplete Universe: Totality, Knowledge, and Truth. Cambridge, MA: The MIT Press.

_____. 1997. "Worlds by Supervenience: Some Further Problems." Analysis 57 (2):146-151.

- Haack, Susan. 1977. "Lewis' Ontological Slum." Review of Metaphysics 30:415-429.
- Hallett, Garth L. 1991. Essentialism: A Wittgensteinian Critique. Albany, NY: SUNY Press.
- Hennig, W. 1979. *Phylogenetic Systematics*. Translated by D. Davis and R. Zangerl. Urbana, IL: University of Illinois Press.
- Hintikka, Jaakko. 1973. Time & Necessity: Studies in Aristotle's Theory of Modality. Oxford: Clarendon Press.

. 1979. "The Modes of Modality." In *The Possible and the Actual: Read*ings in the Metaphysics of Modality, edited by M. J. Loux, 65-79. Ithaca, NY: Cornell University Press.

. 1981. "Aristotle on the Realization of Possibilities in Time." In *Reforging the Great Chain of Being: Studies of the History of Modal Theories*, edited by S. Knuuttila, 57-72. Dordrecht: D. Reidel.

- Hobbes, Thomas. 1839. "De Corpore." In *The English Works of Thomas Hobbes*, edited by W. Molesworth. Aalen: Scientia.
- Hughes, G. E., and M. J. Cresswell. 1996. A New Introduction to Modal Logic. London: Routledge.
- Hume, David. 1975. Enquiries Concerning Human Understanding and Concerning the Principles of Morals. Edited by L.A. Selby-Bigge and P.H. Nidditch. 2d ed. Oxford: Oxford University Press.

_____. 1978. A Treatise of Human Nature. Edited by L.A. Selby-Bigge and P.H. Nidditch. 2d ed. Oxford: Oxford University Press.

- Hymers, Michael. 1991. "Something Less Than Paradise: The Magic of Modal Realism." Australasian Journal of Philosophy 69 (3):251-263.
- Irwin, T. H. 1988. Aristotle's First Principles. Oxford: Clarendon Press.
- Ishiguro, Hide. 1972. Leibniz's Philosophy of Logic and Language. Ithaca, NY: Cornell University Press.
- Jubien, Michael. 1988. "Problems With Possible Worlds." In *Philosophical Analysis*, edited by D. F. Austin, 299-322.

_____. 1993. Ontology, Modality, and the Fallacy of Reference. Cambridge: Cambridge University Press.

- Kant, Immanuel. 1929. Critique of Pure Reason. Translated by Norman Kemp Smith. New York: St. Martin's Press.
- Kaplan, David B. 1994. "A Problem in Possible-World Semantics." In Modality, Morality, and Belief: Essays in Honor of Ruth Barcan Marcus, edited by W. Sinnott-Armstrong, 41-52. New York: Cambridge University Press.
- Khalidi, Muhammad Ali. 1993. "Carving Nature at the Joints." *Philosophy of Science* 60:100-113.

- Kirkham, Richard L. 1992. Theories of Truth: A Critical Introduction. Cambridge, MA: The MIT Press.
- Kitcher, Philip. 1989. "Some Puzzles About Species." In What the Philosophy of Biology Is: Essays Dedicated to David Hull, edited by M. Ruse, 183-208. Dordrecht: Kluwer.
- Kneale, William, and Martha Kneale. 1962. The Development of Logic. Oxford: Clarendon Press.
- Knuuttila, Simo. 1981. "Time and Modality in Scholasticism." In *Reforging the Great Chain of Being: Studies of the History of Modal Theories*, edited by S. Knuuttila, 163-258. Dordrecht: D. Reidel.

. 1982. "Modal Logic." In The Cambridge History of Later Medieval Philosophy: From the Rediscovery of Aristotle to the Disintegration of Scholasticism, 1100-1600, edited by N. Kretzmann, A. Kenny and J. Pinborg, 342-357. Cambridge: Cambridge University Press.

Kripke, Saul A. 1963. "Semantical Considerations on Modal Logic." Acta Philosophica Fennica 16:83-94.

_____. 1971. "Semantical Considerations on Modal Logic." In *Reference and Modality*, edited by L. Linsky, 63-72. Oxford: Oxford University Press.

. 1972. "Naming and Necessity." In *Semantics of Natural Language*, edited by D. Davidson and G. Harman, 252-355. Dordrecht: Reidel.

_____. 1975. "Outline of a Theory of Truth." *Journal of Philosophy* 72:690-716.

_____. 1980. Naming and Necessity. Cambridge, MA.: Harvard University Press.

LaPorte, Joseph. 1996. "Chemical Kind Term Reference and the Discovery of Essence." Nous 30 (1):112-132.

. 1997. "Essential Membership." Philosophy of Science 64:96-112.

Leibniz, G. W. 1989. "On Freedom." In *Philosophical Essays*, 94-98. Indianapolis, IN: Hackett.

Lewis, C. I., and C. H. Langford. 1959. Symbolic Logic. 2d ed. New York: Dover.

Lewis, David K. 1968. "Counterpart Theory and Quantified Modal Logic." Journal of Philosophy 65:113-126.

. 1973. Counterfactuals. Cambridge, MA: Harvard University Press.

_____. 1976. "Survival and Identity." In *The Identities of Persons*, edited by A. O. Rorty, 17-40. Berkeley: University of California Press.

. 1986. On the Plurality of Worlds. Oxford: Blackwell.

- Locke, John. 1975. An Essay Concerning Human Understanding. Edited by P. H. Nidditch. Oxford: Clarendon Press.
- Lopston, Peter. 1996. "Prior, Plantinga, Haecceity, and the Possible." In Logic and Reality: Essays on the Legacy of Arthur Prior, edited by B. J. Copeland, 419-435. Oxford: Clarendon Press.
- Loux, Michael J. 1979. "Introduction: Modality and Metaphysics." In *The Possible* and the Actual: Readings in the Metaphysics of Modality, edited by M. J. Loux, 15-64. Ithaca, NY: Cornell University Press.
- Lovejoy, Arthur O. 1936. The Great Chain of Being: A Study of the History of an Idea. Cambridge, MA.: Harvard University Press.
- Lukasiewicz, Jan. 1957. Aristotle's Syllogistic From the Standpoint of Modern Formal Logic. 2d ed. Oxford: Clarendon Press.
- Lycan, William. 1979. "The Trouble with Possible Worlds." In *The Possible and* the Actual: Readings in the Metaphysics of Modality, edited by M. J. Loux, 274-316. Ithaca, NY: Cornell University Press.

____. 1994. Modality and Meaning. Dordrecht: Kluwer.

- MacColl, Hugh. 1906. Symbolic Logic and its Applications. London: Longmans, Green, and Co.
- Marcus, Ruth Barcan. 1971. "Essential Attribution." Journal of Philosophy 68:187-202.
- Marenbon, John. 1998a. "Boethius: From Antiquity to the Middle Ages." In Medieval Philosophy, edited by J. Marenbon, 11-28. New York: Routledge.

_____. 1998b. "The Twelfth Century." In *Medieval Philosophy*, edited by J. Marenbon, 150-187. New York: Routledge.

Mates, Benson. 1986. The Philosophy of Leibniz: Metaphysics and Language. Oxford: Oxford University Press.

Matthews, Gareth B. 1982. "Accidental Unities." In Language and Logos, edited by M. Schofield and M. Nussbaum, 223-40. Cambridge: Cambridge University Press.

_____. 1990. "Aristotelian Essentialism." Philosophy and Phenomenological Research I (Supplement):251-262.

- Mayr, Ernst. 1970. Populations, Species, and Evolution; an Abridgment of Animal Species and Evolution. 2nd ed. Cambridge, MA: Belknap Press.
- Mayr, Ernst, and Peter D. Ashlock. 1991. Principles of Systematic Zoology. 2nd ed. New York: McGraw-Hill.
- McCall, Storrs. 1963. Aristotle's Modal Syllogisms, Studies in Logic and the Foundations of Mathematics. Amsterdam: North-Holland Pub. Co.
- McGinn, Colin. 1999a. "A Note on the Essence of Natural Kinds." In Knowledge and Reality: Selected Essays, 49-56. Oxford: Clarendon Press.

_____. 1999b. "On the Necessity of Origin." In Knowledge and Reality: Selected Essays, 57-64. Oxford: Clarendon Press.

- McKay, Thomas J. 1986. "Against Constitutional Sufficiency Principles." In *Studies in Essentialism*, edited by P. A. French, T. E. Uehling, Jr. and H. K. Wettstein, 295-303. Minneapolis: University of Minnesota Press.
- McKelvey, Bill. 1982. Organizational Systematics: Taxonomy, Evolution, Classification. Berkeley: University of California Press.
- McMichael, Alan. 1983. "A Problem For Actualism About Possible Worlds." The Philosophical Review XCII (January 1983):49-66.
- Mellor, D. H. 1977. "Natural Kinds." The British Journal for the Philosophy of Science 28:299-312.
- Menzel, Christopher. 1990. "Actualism, Ontological Commitment, and Possible World Semantics." Synthese 85:335-389.
- Mill, John Stuart. 1930. A System of Logic. London: Longmans, Green, and Co.
- Noonan, Harold. 1983. "The Necessity of Origin." Mind 92:1-20.
- Ockham, William of. 1988. "Modal Consequences." In Logic and the Philosophy of Language, edited by N. Kretzmann and E. Stump, 314-336. Cambridge: Cambridge University Press.

Odegard, Douglas. 1974. "Locke's Unnatural Kinds." Analysis 35:208.

- Paterson, H. 1985. "The Recognition Concept of Species." In Species and Speciation, edited by E. S. Vrba, 21-29. Pretoria: Transvaal Museum.
- Patoluoto, Ilkka. 1988. "Hobbe's System of Modalities." In Modern Modalities: Studies of the History of Modal Theories from Medieval Nominalism to Logical Positivism, edited by S. Knuuttila, 71-84. Dordrecht: Kluwer.
- Patterson, Richard. 1995. Aristotle's Modal Logic: Essence and Entailment in the Organon. Cambridge: Cambridge University Press.

Perry, John. 1972. "Can the Self Divide?" Journal of Philosophy 69 (16):463-488.

Plantinga, Alvin. 1974. The Nature of Necessity. Oxford: Clarendon Press.

. 1979. "Actualism and Possible Worlds." In *The Possible and the Actual: Readings in the Metaphysics of Modality*, edited by M. J. Loux, 80-87. Ithaca, NY: Cornell University Press.

Plato. 1997. Plato: Complete Works. Edited by J. M. Cooper. Indianapolis: Hackett.

- Price, M. S. 1982. "On the Non-Necessity of Origin." Canadian Journal of Philosophy 12:33-45.
- Putnam, Hilary. 1975. "The Meaning of 'Meaning'." In Mind, Language, and Reality, 215-271. Cambridge: Cambridge University Press.

. 1990. "Is Water Necessarily H2O?" In Realism with a Human Face, edited by J. Conant, 54-79. Cambridge, MA: Harvard University Press.

_____. 1993. "Meaning and Reference." In *Meaning and Reference*, edited by A. W. Moore, 150-161. Oxford: Oxford University Press.

Quine, W. V. 1953. "Two Dogmas of Empiricism." In From a Logical Point of View: Nine Logico-Philosophical Essays, 20-46. Cambridge, MA: Harvard Univ Press.

_____. 1976a. "On Multiplying Entities." In *The Ways of Paradox and Other Essays*, 259-264. Cambridge, MA: Harvard University Press.

. 1976b. "On Simple Theories of a Complex World." In *The Ways of Paradox and Other Essays*, 253-258. Cambridge, MA: Harvard University Press. _____. 1976c. "Three Grades of Modal Involvement." In *The Ways of Paradox* and Other Essays, 158-176. Cambridge, MA: Harvard University Press.

- Rescher, Nicholas. 1979. "The Ontology of the Possible." In The Possible and the Actual: Readings in the Metaphysics of Modality, edited by M. J. Loux, 166-181. Ithaca, NY: Cornell University Press.
- Richards, T. O. M. 1975. "The Worlds of David Lewis." Australasian Journal of Philosophy 53:105-118.
- Russell, Bertrand. 1908. "Mathematical Logic as Based on the Theory of Types." American Journal of Mathematics 30:222-262.

_____. 1996. The Principles of Mathematics. New York: Norton.

- Sainsbury, R. M. 1995. Paradoxes. 2d ed. Cambridge: Cambridge University Press.
- Salmon, Nathan U. 1981. *Reference and Essence*. Princeton, N.J.: Princeton University Press.
- Scotus, John Duns. 1994. Contingency and Freedom, Lectura I 39. Translated by A. Vos Jaczn, H. Veldhuis, A. H. Looman-Graaskamp, E. Dekker, and N. W. Den Bok. Edited by N. Kretzmann. Dordrecht: Kluwer.
- Serene, Eileen F. 1981. "Anselm's Modal Conceptions." In Reforging the Great Chain of Being: Studies of the History of Modal Theories, edited by S. Knuuttila, 117-162. Dordrecht: D. Reidel.
- Shalkowski, Scott A. 1994. "The Ontological Ground of the Alethic Modality." The *Philosophical Review* 103 (4):669-688.
- Skyrms, Brian. 1976. "Possible Worlds, Physics and Metaphysics." *Philosophical Studies* 30:323-332.
- Sober, Elliot. 1993. Philosophy of Biology, Dimensions of Philosophy Series. Boulder, CO: Westview Press.

Stalnaker, Robert. 1976. "Possible Worlds." Nous 10:65-75.

. 1977. "Complex Predicates." The Monist 60 (3):327-339.

Tarski, Alfred. 1983. Logic, Semantics, Metamathematics. Translated by J.H. Woodger. 2d ed. Indianapolis: Hackett.

- Tomberlin, James E., and Frank McGuinness. 1994. "Troubles With Actualism." In *Philosophical Perspectives*, 8: Logic and Language, edited by J. E. Tomberlin, 459-466. Atascadero: Ridgeview.
- Uzgalis, William L. 1988. "The Anti-Essential Locke and Natural Kinds." *Philosophical Quarterly* 38:330-338.
- van Inwagen, Peter. 1985. "Plantinga on Trans-World Identity." In Alvin Plantinga: A Profile, edited by J. Tomberlin and P. van Inwagen, 101-120. Dordrecht: D. Reidel.
- Van Valen, L. 1976. "Ecological Species, Multispecies, and Oaks." Taxon 25:233-239.
- Vienne, Jean-Michel. 1993. "Locke on Real Essence and Internal Constitution." Proceedings of the Aristotelian Society 93:139-153.
- Waterlow, Sarah. 1992. Passage and Possibility: A Study of Aristotle's Modal Concepts. Oxford: Clarendon Press.
- Whitehead, A. N., and B. A. W. Russell. 1910. Principia Mathematica. 1st ed. 3 vols. Cambridge: Cambridge University Press.
- Wiggins, David. 1980. Sameness and Substance. Oxford: Basil Blackwell.
- Wilkerson, T.E. 1995. Natural Kinds. Aldershot: Avebury.
- Wittgenstein, Ludwig. 1922. Tractatus Logico-Philosophicus. Translated by C.K. Ogden. New York: Routledge.
- Wright, Crispin. 1986. "Inventing Logical Necessity." In Language, Mind and Logic, edited by J. Butterfield, 187-209. Cambridge: Cambridge University Press.