Topics in the Philosophy of Possible Worlds

Daniel Patrick Nolan

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University.

Submitted for examination in February 1997.
Except where otherwise indicated in the text, this thesis is entirely my own work.

Daniel Nolan
Acknowledgements

For My Family

I have been fortunate in being able to discuss many of the ideas in this thesis with many people who had interesting and useful things to say (though they should not be considered to blame for the views I will express). To my supervisors, Frank Jackson and Peter Menzies, I owe the obvious debt of gratitude. As do the other people whom I have been at one time or another on my supervision panel—Graham Oddy, Phillip Pesch and Doug Samuel. I also have benefited from discussions with other staff, students and colleagues at the ANU, especially David Armstrong, James Chalm, Steve Chisholm, Jenny Howelsthorne, Cathy Legg and Timothy Williamson, but of course many others. I have also received much useful feedback from audiences at conferences and departmental seminars, and I thank all those who have contributed.

The copyright for substantial sections of Chapter Four and Chapter Six belongs, or will soon belong, to Kluwer Academic Publishers.
Acknowledgments

I have been fortunate in being able to discuss many of the issues in this thesis with many people who had interesting and useful things to say (though they should not be considered to blame for the views I will express). To my supervisors, Frank Jackson and Peter Menzies, I owe the obvious debt of gratitude, as I do to the other people who have been at one time or another on my supervision panel: Graham Oppy, Philip Pettit and Greg Restall. I also have benefited from discussions with other staff, students and visitors to the ANU, especially David Armstrong, James Chase, John O’Leary-Hawthorne, Cathy Legg and Timothy Williamson, but of course many others. I have also received much useful feedback from audiences at conferences and departmental seminars, and I thank all those who have contributed.

The copyright for substantial sections of Chapter Four and Chapter Six belongs, or will soon belong, to Kluwer Academic Publishers.
Abstract

This thesis deals with several issues in the philosophy of possible worlds. In the first two chapters several conceptual issues are addressed. In the first chapter two senses of "possible world" are distinguished, and an account is offered of what it would take for entities which were putatively the possible worlds to in fact count as possible worlds. The second chapter is concerned with the issue of what counts as a modal primitive, and the general question of what it is for something to be a theoretical primitive. Chapter Three deals with the Hazen cases, which are a well-known set of ordinary-language modal claims which have possible-worlds translations but supposedly do not have modal-operator translations. I show that they can be represented in modal-operator languages, and that in some respects it is more natural to do so. In Chapter Four I outline four major problems for a position I call strong modal fictionalism, a theory of possible worlds gaining increasing popularity. Chapter Five provides a solution to the problem of how certain forms of abstractionism and fictionalism about possible worlds can allow for the possibility of alien properties and relations. Chapter Six contains an argument that an account of possible worlds should contain an unrestricted version of what is known as the "principle of recombination", and examines what follows if this unrestricted version of the principle is employed, and finally Chapter Seven shows that, if there are enough possibilia, and some assumptions are made about the respectability of mereology and plural quantification, it becomes possible to provide an interpretation of class theory that requires neither any distinctively mathematical ontology nor any distinctively mathematical ideology. Appendix One contains proofs that systems discussed in Chapter Seven provide interpretations of class theory which satisfy a standard set of class axioms, and Appendix Two discusses a problem for an argument David Lewis provides to support his conception of classes.
INTRODUCTION ......................................................................................................................... 1

THE GENERAL APPROACH ........................................................................................................ 1

CHAPTER ONE: POSSIBLE WORLDS ......................................................................................... 5

1. WHAT ARE “POSSIBLE WORLDS”? .................................................................................. 5

S2. THE COMPETING THEORIES OF POSSIBLE WORLDS ......................................................... 20

CHAPTER 2 — TOWARDS EXPLAINING THE CONNECTION BETWEEN MODALITY AND POSSIBLE WORLDS ................................................................................................. 28

1. WHAT MIGHT THE CONNECTION BETWEEN POSSIBLE WORLDS TALK AND MODAL OPERATOR TALK BE? ........................................................................................................... 29

2. THE PRIMITIVES OF A THEORY .......................................................................................... 33

3. MODAL PRIMITIVES ............................................................................................................. 42

CHAPTER THREE — EXPRESSIVE COMPLETENESS WITHOUT POSSIBLE WORLDS: THE HAZEN CASES .............................................................................................................................. 51

INTRODUCTION .......................................................................................................................... 51

1. NEUTRAL QUANTIFICATION .................................................................................................. 52

2. THE FIRST THREE HAZEN CASES: PROBLEMS OF EXISTENCE ........................................... 55

3. HAZEN CASE NUMBER FOUR: A CASE TO BE SET ASIDE .................................................. 60

4. TWO CHALLENGING HAZEN CASES .................................................................................... 62

5. DEFENDING THE USE OF SECOND-ORDER QML .................................................................. 68

CHAPTER 4 — PROBLEMS FOR ‘STRONG’ MODAL FICTIONALISM ........................................... 77

INTRODUCTION .......................................................................................................................... 77

1. THREE VARIETIES OF MODAL FICTIONALISM: BROAD, TIMID AND STRONG ....................... 79

2. THE FIRST PROBLEM: ARTIFICIALITY .................................................................................. 81

3. THE SECOND PROBLEM: THE INCOMPLETENESS OF THE MODAL FICTION ................................ 87

4. THE THIRD PROBLEM: PROPOSITIONS .................................................................................. 89

5. AN ASSUMPTION RELIED ON BY THE INCOMPLETENESS OBJECTIONS, AND A DEFENCE THEREOF ................................................................................................................................. 93

6. A FINAL PROBLEM: REDUNDANCY ....................................................................................... 93

7. CONCLUSION .......................................................................................................................... 94

CHAPTER 5 — REPRESENTING ALIENS ....................................................................................... 95

1. THE PROBLEM .......................................................................................................................... 96

2. DENIALS THAT THE PROBLEM ARISES ................................................................................ 98

3. EXPLAINING THE CAVEAT “NOT CONSTRUCTIBLE OUT OF EXISTING UNIVERSALS” ................. 100
4. HOW TO REPRESENT ALIEN UNIVERSALS .................................................. 104
5. THE PROBLEM FOR LINGUISTIC ERSATZISM AND MODAL FICTIONALISM ................................................. 107
6. MAKING SURE THE WORLD-BOOK REPRESENTS THE ACTUAL UNIVERSE ........................................... 114
7. FROM UNIVERSES TO WORLDS — SALVAGING ERSATZ WORLDS FROM THE INTER-
   DEFINITION .................................................................................................................. 119
8. HOW A LINGUISTIC ERSATZER CAN REPLY TO LEWIS'S OBJECTION BASED ON ALIEN
   UNIVERSALS ........................................................................................................... 125
9. CONCLUSION ........................................................................................................... 127

CHAPTER 6 — RECOMBINATION UNBOUND ........................................ 128

1. THE PRINCIPLE OF RECOMBINATION AND LEWIS'S RESTRICTION .................................................. 128
2. THE FORREST/ARMSTRONG ARGUMENT ................................................................................ 130
3. WHY POSSIBILIA SHOULD BE SEEN AS FORMING A PROPER CLASS ........................................... 136
4. SOME ARGUMENTS AGAINST THE THESIS THAT THERE IS A MAXIMUM SIZE OF WORLDS (MS) ........... 143
5. CONCLUSION ........................................................................................................... 146

CHAPTER 7 — A REDUCTION OF CLASSES TO POSSIBLE WORLDS ....... 147

1. HOW TO REDUCE CLASSES TO POSSIBLE WORLDS ...................................................................... 147
2. HOW ATOMS CAN PERFORM DOUBLE DUTY .................................................................................. 153
3. A SURPRISING FEATURE OF THE ACCOUNT .................................................................................. 159
4. UTILISING THE SURPRISING FEATURE ......................................................................................... 163

RECAPITULATION ................................................................................................................. 169

APPENDIX ONE: PROOFS FOR CHAPTER 7 ................................................................. 172

SECTION 1: PROVING THAT MY SYSTEMS SATISFY LEWIS'S SO-CALLED "STANDARD AXIOMS" .............................. 172
SECTION 2: PROOFS OF IMPREDICATIVE COMPREHENSION ........................................................................... 187

APPENDIX 2: TROUBLE FOR LEWIS'S "MAIN THESIS" ......................................................... 189

I. THE PROBLEM ........................................................................................................... 189
II. THE MODEL .............................................................................................................. 191
III. THE REMEDY .......................................................................................................... 193

REFERENCES ...................................................................................................................... 195
Introduction

I began work on this thesis hoping to provide a satisfactory and correct account of both the nature and function of modality, and its relation to possible worlds — in short, to solve all of the outstanding problems in the philosophy of modality, including the questions about the existence and nature of possible worlds. Unsurprisingly, a project which has been pursued for decades by many of the important figures of twentieth century philosophy was not one which I was able to complete in a few dozen months. Instead, I have settled for chipping away at various of the challenges and exploring some of the options available, and I will be happy if I have been able to make some progress in a vast field in which so much has been done and in which so much remains to be done. I will not purport to produce something that could be the final word on the topic of the nature of modality and possible worlds and their connections with other important issues — but progress in such areas is not an unworthy goal, and I have some hope that I have achieved it.

In each of the chapters of this thesis I will address a problem in the philosophy of possible worlds. In this introduction, however, I will seek to put these somewhat isolated projects in the context of an overall approach to modality and possible worlds. The chapters do not rely on features of this general approach — even if my overall attitudes to this area were completely misguided I hope that the chapters could still stand alone as contributions to the discussion of the issues they are concerned with. There is a certain amount of independence in the other direction as well: the general approach I favour could still be followed even if the conclusions reached in the following chapters were incorrect. Fortunately, however, there does not need to be necessary dependence in either direction for a general strategy to help motivate specific investigations, nor for the outcome of specific investigations to lend support to an overall approach. The chapters of this work are only fragmentary parts of what I hope will become a more fully developed research program, but I hope that they are a good start. I do not here intend to provide a general defence of the views I will set out as constituting my general philosophical position: general defences of each of many of these views would be theses in themselves, and I have chosen to write this thesis rather than any of them. The next section of this introduction, then, will outline my general approach and explain how the chapters to follow are to be seen in the context of that approach.

The General Approach

I am a metaphysical actualist about modality, in at least two senses. The first is that I think that the modal truths (or facts, or what-have-you) are true in virtue of the universe
we live in — the "actual world"\(^1\) — or sections thereof, at least. This is a controversial position, but one shared by many. The second way in which I am an actualist is more radical — I suspect (believe/hope/wish) that the modal truths, and the truths about possible worlds, can be perspicuously explained in terms of things or states of affairs in the actual world that can be characterised in non-modal terms — some interesting "reduction" of the modal into the non-modal is available. This is a view shared by few, and I disagree with most of them. I do not agree with any of the plethora of idealists, conventionalists or constructivists who are among the most prominent holders of such a view.

As well as this actualism, there is another important prejudice or hunch that I find myself in sympathy with. It is that talk about possible worlds, and possible worlds themselves if there are such things, is/are to be explained in terms of modal truths, or modal facts (able to be cast in the traditional language of the modal operators, or modally charged predicates such as predications of essentiality, or conditional language or the language of entailments, and so on). In holding this view, I set myself at odds with a strong tradition in contemporary philosophy which seeks (or takes for granted) an explanation of such modal discourse in terms of possible worlds, and seeks the grounds for the truth of modal claims in an ontology of possible worlds. On the other hand, my position is not in the mainstream of views which seek to analyse possible worlds in terms of modality, or at least refuse to analyse modality in terms of possible worlds, as many such views are primitivist about modality, whereas I hope for a further reduction.

Despite my view that possible worlds are not useful for providing the grounds of modality, or providing a metaphysically interesting explanation of modality, I am not among those that would eschew possible worlds altogether. The worth of talking about possible worlds has been proved again and again in facilitating metaphysical investigations, in providing attractive frameworks for accounting for intentional states and propositions, for providing a better heuristic grasp of counterfactuals, in assisting in providing a model-theoretic framework for evaluating modal inferences, and so on. I think talk about possible worlds is very useful, and I want to salvage the truth, or at the very least the usefulness, of many of the assertions that are made in terms of possible worlds (or possible situations, or possible outcomes, or alternatives (where some of those are non-actual), or whatever). I think the use of possible worlds talk extends even to reasoning about modal issues — though I see their value in such cases as heuristic, and sometimes talk of possible worlds is no more than convenient shorthand. So the final plank in my platform (such as it is) is that possible worlds should be salvaged and explained, despite their not having the fundamental role in explaining matters modal, and

\(^1\) In this thesis, I will be employing quotation marks not merely to signal the use-mention distinction, but for the many other purposes for which they are employed in English as well. Context should
Despite the prospect that all such matters may be reducible to components of the actual world.

The first two chapters of this thesis will be primarily concerned with issues of conceptual analysis. The first chapter deals with a cluster of conceptual issues which are important for an understanding of what I and others might be talking about when we discuss "possible worlds" and theories thereof. I shall discuss what is meant by the expression "possible world", and argue that it is ambiguous in a way that can be confusing; what the adequacy conditions on identifying things as possible worlds might be; and providing a discussion of the main ways people in fact have accounted for "possible worlds". The second chapter will address an important set of conceptual issues which do not seem to have been treated in a way I find entirely satisfactory in the literature. It deals with the question of what it might be to provide satisfactory explanations of modality and possible worlds, and if one is to be explained in terms of the other what sort of theoretical relationship they must have. It deals with the issue of how in general we are to evaluate the cost of the rival theories of modality, when we should treat theoretical primitives as "modal primitives", and what it could be to provide an account of the modal in terms of the non-modal. Chapter 2 does not attempt to evaluate or compare specific theories, let alone to construct a theory which explain the modal in terms of the non-modal, even though I hope that this might be able to be done. Rather, it is a discussion of some prior questions concerning how one could evaluate and compare the theoretical resources of modal theories, and what it would take for an account to be an account of the modal in terms of the non-modal.

Chapter 3 of this thesis deals with a threat to the idea that talk about possible worlds is to be analysed in terms of modal discourse (and the corresponding metaphysical claim that possible worlds are to be analysed in terms of modal facts, or otherwise in terms of modality). This threat consists of examples of claims which, according to those who favour analysis of modal discourse in terms of possible worlds rather than vice versa, show that ordinary language has the resources to express things that ordinary formal systems employing modal operators are unable to represent, but which formal systems employing quantification over possible worlds are able to represent. If it were true that there were apparently modal claims formulable in ordinary language which could be represented in terms of quantification over possible worlds but which resisted formulation in a language which had as its modal resources instead only the standard modal operators, this would be a prima facie reason for preferring quantification over possible worlds to the employment of unreducible modal operators. As I will argue, however, the ordinary language modal claims which I will discuss that supposedly cannot be represented with modal operators can in fact be so represented, and furthermore there is a general reason, for at least one class of such controversial cases, for supposing that disambiguate which function they are to serve in any particular case.
quantification over possible worlds is better seen as an unnecessarily restrictive variant of a more general approach which employs modal operators rather than the device of quantification over possible worlds.

Chapters 4 and 5 discuss problems facing certain accounts of possible worlds which respect actualism, in the sense that they do not take anything more to exist than the actual world and/or its contents. In Chapter 4 I criticise the pretensions one approach might have to provide a reductive account of modality in terms of possible worlds — that variety of account of possible worlds talk known as “modal fictionalism”, which has been having an increase in popularity in previous years, if the volume of discussion in the journals is anything to go by. Chapter 5, on the other hand, deals with a problem involving “alien” universals, which is meant to cause trouble for both modal fictionalism and some other varieties of actualism concerning possible worlds. I show that this problem has a solution, and so to that extent defend some actualist accounts of possible worlds against what has been thought to be a serious objection. These chapters do not between them constitute anything like an exhaustive discussion of even the most interesting and important issues facing actualist theories of possible worlds, and nor are they intended to. Through an attack on an increasingly popular version of a “possible worlds first” actualist account of possible worlds, and a defence of some tempting actualist accounts against an objection which has been raised against them respectively, these two chapters do (at least potentially) serve the ends of the wider project.

In Chapter 6, I examine a variety of cardinality argument based on a plausible principle governing possibilia known as the “principle of recombination”. The cardinality objection against some varieties of theories of possible worlds incorporating this principle is known as the “Forrest/Armstrong cardinality argument”, after Peter Forrest and D.M. Armstrong who formulated it. Chapter 6 discusses this argument, and a related cardinality argument based on the principle of recombination, and comes to the conclusion that one ought accept a proper class of possibilia in light of these considerations. Finally, in Chapter 7, I examine a benefit available if one adopts the approach to the cardinality of possibilia defended in Chapter 6: it becomes possible to reduce class theory to a theory of possible worlds. Chapter 7 in part carries on the discussion of Chapter 6 on the impact of cardinality arguments of various sorts, but is also a useful discussion of one way of applying a theory of possible worlds to provide answers to philosophical questions: in this case, they can provide the ontology of mathematics.

A substantial part of chapter 4 is due to be published as “Three Problems for ‘Strong’ Modal Fictionalism” Philosophical Studies 1997, and a slight variant of chapter 6 is due to be published as “Recombination Unbound” Philosophical Studies 1996. The copyright of much of the contents of the two chapters will therefore belong to Kluwer Academic Publishers.
Chapter One: Possible Worlds

1. What are "Possible Worlds"?

Since my thesis is concerned with issues in the philosophy of possible worlds, and since there is much debate about "what possible worlds are", I feel I should say something about what the expression "possible worlds" means, at least in my mouth and hopefully in the literature as well. One of the problems with "possible worlds" is that that expression is ambiguous, and this may lead to some confusion, although most people writing in this area are aware of the ambiguity, and so the negative effects of ambiguity are limited. What follows then, is some conceptual analysis of the expression "possible world", where I distinguish two rival senses, and make explicit the most basic adequacy conditions on a theory which purports to explain the nature of possible worlds.

Conceptual analysis is a tricky matter: Wittgenstein's dictum "If one tried to advance theses in philosophy, it would never be possible to debate them, because everyone would agree with them" (Wittgenstein 1958 s128) is certainly incorrect as a claim about philosophy, but it is more accurate when adapted as a claim about conceptual analysis — for if a piece of conceptual analysis is not agreed to by all of those who are competent in their deployment of the concepts, then this is ipso facto evidence that it is incorrect. So I should not hope my conceptual analysis will be controversial. On the other hand, I hardly wish to bore the reader with banality either — so I must hope for an analysis which will be illuminating through making the implicit explicit rather than through dispelling error (except for error born of confusion).

1.1 Two Conceptions of Possible Worlds

The first conception of possible worlds which I will discuss may, in a sense, be more "commonsensical". My guess is that it is the conception that those who employ pretheoretic talk about possible situations often use, and is the conception used (perhaps unconsciously) by many philosophers, scientists etc. when they are talking about issues involving modal questions (though probably less so when they discuss modality itself). It takes the talk of possible worlds (or possible situations, or possible occurrences or things) very much at face value. According to this conception, a possible state of affairs is a state of affairs, a possible world is a world, a possible talking donkey is a talking donkey, a possible bush fire is a bush fire, and so on. This approach might be labelled the "adjectival" approach, in that it takes modal terms that are syntactically adjectival (such as "possible" in the above examples) to be semantically adjectival as well in that, for example, they can be dropped from an appropriate class of true statements and preserve
truth. For example "a red ball is coloured" only if a ball is coloured, on the straightforward reading of the claim, and "a sad king is rich" is true only if a king is rich (again, on the same reading). Contrast, for example, "non-adjectival" constructions: a prefix like "non-": "a non-king is rich" might be true even if "a king is rich" is not. (The prefix is of course not syntactically an adjective, but I am sure the reader gets the idea.) Other examples of non-semantically adjectival terms include such words as 'putative' and 'fake': the putative king might be no king at all, and the fake king certainly is not.

It is this sort of conception which explains some of the hostility to possible worlds: if there is only one world, only one universe, or whatever, then all this talk about possible worlds where swans are blue or the laws of nature are different is just false. Of course, to have this conception of possible worlds is not the same as saying that they do not exist: most famously, David Lewis has just this conception and nevertheless assures us that possible worlds exist. Of course, he provides us with various explanations (e.g. in terms of tacitly restricted quantifiers) to explain why it is that we think there are no unicorns or talking donkeys, when in fact there being an infinite number of possibilities in which donkeys talk or unicorns exist mean that there are infinite numbers of both unicorns and talking donkeys, according to Lewis. These explanations have some plausibility too: I don't think any "ordinary language argument" directed against Lewisian realism is powerful enough to reject it out of hand.

In fact, as Lewis sometimes argues (and I think rightly), there are ordinary language intuitions which support this conception of possible worlds. If there is a possible world where I win an election, then surely the possible world must contain me and an election for me to win! How does a possible world that contains only linguistic entities or mysterious simples manage to be something that I am in, or be the sort of thing in which something mundane like an election can exist? Swans (even possible blue swans) would swim in lakes rather than abstracta, surely! Possible events and processes are often said to occur in possibilities, or ways things could be, or possible states of affairs, and if such things (which are all meant to be, more-or-less, possible worlds or parts of possible worlds) do contain events and processes, then it seems they are things like the actual world and the various actual states of affairs which go to make it up (or occur in it, or are contained in it, or whatever). When it is remembered that the coming into existence of things, as well as their persistence and their ceasing to be are also events or processes, the picture of possible worlds as other things just like the actual can become quite seductive. This attractiveness is one (though certainly not the only one) of the features, curiously, of both Lewisian realism and eliminativism about possible worlds.

The second conception of possible worlds is more likely to be found in discussions specifically about the ontology of possible worlds, as well as among those people who approach modality through model-theoretic mathematical devices. According to this view, possible worlds are not necessarily meant to be much like the great region of
space-time that we all find ourselves in, nor are *possibilia* much like tables or chairs or electrons or galaxies. Instead, possible worlds might be made up of propositions, or are perhaps set-theoretic constructions, or maybe highly abstract entities, or even perhaps unusual, almost supernatural entities connected to everyday things by specialised and otherwise quite exotic relations. Or perhaps, depending which specific account is being relied on, they may be meant to be something quite different again. Such views are often motivated by the thought that merely possible entities and merely possible worlds are quite different sorts of things from the actual: it is plausible to suppose that whatever it is about a possible world that makes it true that there are possible talking donkeys “in” it, it is not that there really are talking donkeys roaming around inside it (there is no such thing as a talking donkey, after all).

If the concept of possible worlds is not to be explicated in terms of their being the same sort of thing as the world we are all parts of, then what is it, according to this conception, for something to be a possible world? Part of the answer to this is, of course, that something is true according to some possible world or other if that claim is possibly true, and a claim is necessarily true if it is true at all possible worlds. Of course, this does not explain what the relation of "true according to" is, but this does not make the correspondence condition objectionably mysterious, for a correspondence condition of this sort can be stated to provide a condition which any pair of <objects purported to be possible worlds, a relation purported to be a "true according to" relation> must satisfy if (but perhaps not only if) they are to be satisfactory candidates for the roles. The claim about possible worlds and the "true according to" relation is as follows:

possibly P iff P has the "true according to" relation to some possible world, and necessarily P iff P has the "true according to" relation to all possible worlds (where P is any proposition).

It seems to me a conceptual truth that this schemata must hold for anything that is to count as a possible world: I am tempted to say that were there to be candidates for the name that were as satisfactory as one could wish in all other areas, yet they would not be possible worlds if they did not satisfy the above condition (or rather if they plus their candidate "true according to" relation would not jointly satisfy the condition). However, I resist such temptation, as conceptual claims (especially philosophically interesting ones) are almost never that crucial — what I am prepared to assert is that the above scheme must be mostly satisfied by any candidate (or, more precisely, candidate pair) for it to be

---

2 I am using "proposition" here in a quite general sense, to mean the things that are most apt for assessment as true or false in a linguistic sense. I leave it to philosophy of language to fill in the details of what exactly propositions are.
a successful one. I cannot really offer too much argument for taking this condition to be so important however: as with most claims about the importance of one concept to another, either one sees the truth of it immediately or one does not.

Is the satisfaction of this condition a sufficient condition for a group of objects to be possible worlds? (or again, to be more strict, a group of objects and a relation to be possible worlds and the "true according to" relation)? I suspect not, although I would have a hard time trying to budge someone who insisted that this is all there is to it. Some comments, though, for those not yet stuck in this insistence:

The first comment is that this leaves open the possibility (at least the doxastic possibility) that more than one group of objects (object-relation pairs) might have claims to be possible worlds. Not that this is necessarily a problem — there may well be more than one set of objects that each deserve the name in their own way3. Jadeite and Nephrite both have adequate claims to be called Jade, and something similar may be true of different types of possible world. The analogy is not exact, since each system of possible worlds needs to be exhaustive in a certain sense: if a group (call it A) and its relation are adequate deservers of the title of possible worlds, then when we quantify over all possible worlds we want to be quantifying only over members of A. We may also be able to do the same trick with another group (B) — but in the context of treating B as the possible worlds, we must take all the possible worlds to be members of B. So perhaps we may have to say that neither A nor B deserve the title "possible worlds" simpliciter — we may have to complicate our account so that being a possible world is something that can happen only in the context of an interpretation (an object 'b' might be a possible world in the context of treating B as the group of possible worlds, but not a possible world in the context of treating A as the group of possible worlds). So taking the correspondence condition to be a sufficient condition (in one sense) for something to be a possible world leads to potential multiple-realiser problems.

In fact, this problem is worse than it may appear. For if it is granted that any class of ordered pairs can be used to define a two-place relation, then any group of objects can satisfy the correspondence condition (provided there are enough of them), if any can. For take some ordered pair of <a group of objects, a relation> that satisfies the correspondence condition. Call this group W and its "true according to" relation R. Each member of W will be R related to certain propositions and not others (in fact each W will be presumably R related to an infinite number of propositions). Now, take any other group of objects of sufficient size (call it A). Pair the members of A in a one-one fashion

3 Compare, for example, a Lewisian who admits (as Lewis does) that certain set-theoretic constructions are perfectly adequate "possible worlds" for some purposes, even though the parallel concrete worlds are also perfectly adequate possible worlds.
with members of \( W \) (it doesn't really matter how)\(^4\). Then construct a set of ordered pairs so that for each member of \( W \) and each of the propositions it is \( R \) related to, it includes all and only the ordered pairs of the member of \( A \) which corresponds to that \( W \) and each of those propositions (in the order \(<\text{member of } A, \text{proposition}>, \text{of course} \). This set will then be isomorphic to the extension of \( R \), except that instead of members of \( W \) it will contain the corresponding members of \( A \). The relation defined by this set plus the group \( A \) will jointly satisfy the correspondence condition.

Even assuming that there is only one group \( W \) large enough to satisfy the correspondence condition will not help (in addition to its being wildly implausible). For there will still be all the relations which correspond to the one-one mappings of \( W \) onto itself which do not yield identity. At the very least, there will be an infinite number of such relations.

It should surely give one pause to realise that any large enough group of objects can satisfy the correspondence condition when one uses the right relation. Surely one would not want to say that my coffee cup is a possible world where goblins hunt talking donkeys, but there will be a way of making this come out correct if the correspondence condition is the only restriction on things being possible worlds.

Things look grim for the claim that the correspondence condition is "all we know, and all we need to know" about the conditions for a group to be a group of possible worlds. One could of course deny that every set of ordered pairs defines a relation, but since "being the first and second members respectively of an ordered pair which is a member of \( S \)" where \( S \) is a specific set looks like it expresses a relation (in some sense of relation), this is a hard row to hoe. One approach to solving the problem of what possible worlds are which seems similar to the "correspondence as a sufficient condition" approach, but does not suffer its many-realiser problem or its counterintuitive results as to what various possible worlds are is a "structuralist approach" to the problem, which would be somewhat similar to the structural approach taken to the set-membership relation in the appendix of Lewis's *Parts of Classes* (Burgess, Hazen and Lewis 1991).\(^5\)

This would be to take the correspondence condition as sufficient for defining a certain class of pairs of groups and relations, but to refrain from identifying any of the group-relation pairs as being a pair of the group of possible worlds and the relation of being "true according to". Rather, one could take the truth condition of claims about possible worlds and the "true according to" relation as being given by the truth or falsity of

---

\(^4\) Resources as strong as the Axiom of Choice (or Global Choice, if there should turn out to be a proper class of worlds — see chapters Six and Seven) will ensure that the necessary one-one correlations exist. Without them we do not in general have any such guarantee, but may still be able to do this for some specifications of \( A \), depending on what functions there happen to be.

\(^5\) For a more detailed discussion of this structuralist approach, see Chapter Seven.
appropriate universal quantifications over the group-relation pairs. A claim about possible worlds would be true iff it would be true of all the groups (or all the members of the groups, depending on whether the claim was collective or distributive), and false otherwise. For instance, the claim "there is a possible world according to which Madonna becomes president of the USA" is evaluated by seeing whether every group-relation pair that satisfies the correspondence condition is such that a member of the group is related by the relation to the proposition that Madonna will become president of the USA. If they all are, then the claim is true — if not, then the claim is false.

Such structuralism would seem to do the job we require from possible worlds — and perhaps it is the way to go (or at least one of more acceptable rival ways to go). Perhaps all we need for theoretical purposes to ask of the scheme for determining the truth of claims about possible worlds is that they obey the correspondence condition, never mind how artificial their conformity is. It is unlikely that this would be attractive to someone who thought that possible worlds provided the materials for an explanation of modality (rather than, for instance the other way around), but even they might be able to accept the structuralist construal, especially if they were temperamentally pragmatists. But structuralism is probably a revision of the second conception of possible worlds, albeit an interesting one. It seems to me that there are another two components to the second conception (though since I do not hold that this is completely self-evident, so it seems I run the grave risk of being merely stipulative here). Nevertheless, I will put them forward and hope thereby to elicit sympathetic intuitions from the reader.

The other two components of the second conception of possible worlds are that as well as satisfying the correspondence condition, possible worlds will i) somehow form a reasonably natural kind and ii) have something perspicuously to do with modality. Not any old group with any old gerrymandered relation will do. If this is genuinely part of the conception of possible worlds, then what people should do were they to discover that there were no non-gerrymandered group-relation pairs that satisfied the correspondence condition would be to conclude that possible worlds did not exist. My bet is that this is what would happen. And if there are some deservers of the title "possible world" that do share some natural kind, and in addition their "true according to" relation is a natural non-gerrymandered relation, then general considerations about reference and naturalness would suggest that this group-relation pair has a privileged claim to the title.6 Ensuring naturalness is a general problem, and so one of the general solutions as to how to do this could be adapted to select between all the candidates that satisfy the correspondence

---

6 For a discussion of what naturalness might have to do with eligibility to be the referent of a term, see O'Leary-Hawthorne 1984. I agree with O'Leary-Hawthorne that naturalness should be taken to be a prima facie constraint on how well a given candidate deserves to be the referent (or extension, or whatever) of a term even if it is not one of the 'platitudes' that the referent (or extension) be a natural one. My use of the phrase "part of the conception of possible worlds" should be read in this light.
condition. We might require a certain degree of naturalness in the sets (or classes) of the group and the relation (a la Lewis 1983), or we might require that the relation does not supervene too messily on underlying universals and that the group has some universal in common (or some property in common which is grounded on a not-too-messy supervenience base of universals) a la an Armstrongian approach (see Armstrong 1989), or that the set containing the group members and the set defining the relation be theoretically "cosy" (a la Taylor 1993) etc. Any of these requirements added as an additional requirement to the correspondence condition would serve to filter out some undesirable candidates.

So much for a 'naturalness' condition. What about condition (ii) — that the group-relation pair should have some special relation to modality over and above merely satisfying the extensional requirement that the correspondence condition imposes? For a start, one might wonder why this requirement is needed. Again, since this is meant to be a matter of conceptual analysis I do not have much to say to those who do not find this intuitive, but it seems plausible to me at least that possible worlds, as usually conceived, do not obey the correspondence condition through sheer fluke. It is conceptually coherent to suppose that there could be a group-relation pair (even a natural group-relation pair) that was not a <possible worlds, "true according to" relation> pair, even though it satisfied the correspondence condition, if that satisfaction were a sheer fluke. Consider, for example, what we would conclude if we found that the stars in another galaxy, when viewed from a certain angle, spelled out in English totally detailed descriptions of every possible universe (leave aside size problems and expressibility restrictions for the moment). Would the stars in that galaxy then be what made up possible worlds? I am hoping that our intuitions would suggest that they would not. Whatever possible worlds might be, it is counterintuitive to suppose that they could be made up entirely of actual balls of flaming gas.

How might we capture this intuition that the conformity with the correspondence condition must not be sheer fluke? One way to do this that seems attractive is to insist that there be a specification by means of which the group-relation pair can be identified which does not itself state that the correspondence condition is satisfied, and furthermore that this specification is independently motivated (say, through the objects and relation involved having some other salient connection with modal truths). This is not to say that the correspondence condition is not a vital one for a group of putative possible worlds to be adequate for the name, but rather that a group that deserves the name will satisfy the correspondence condition in virtue of an underlying connection that they have with modality. Requiring that there be a salient connection of worlds to modality in virtue of which the correspondence condition is satisfied is what I take to be the substance of this

---

7 This thought experiment is a variation of one of David Lewis's (Lewis 1986c p 180)
last restriction on candidates for the status of possible worlds.

To recapitulate the positive side of the second conception of possible worlds then: according to this conception, possible worlds 1) obey the correspondence condition, 2) form a non-gerrymandered kind and have a non-gerrymandered "true according to" relation, and 3) have an independent and salient connection with modality which grounds their conformity to the correspondence condition. Note that the third condition would, if accepted, seem to rule out structuralism of the sort I have described, as the class of group-relation pairs that the structuralist analysis uses to evaluate claims about possible worlds was defined in terms of satisfaction of the correspondence condition and not in terms of some independent connection with modality. No-one need be worried by this, not even structuralists about modality of the sort described. All this observation means (if correct) is that they have a revisionary concept of possible worlds.

This second conception of possible worlds is in several ways more abstract: it is a conception of possible worlds in terms of broad conditions they must satisfy and general truths about them, rather than the more specific specification offered by the first conception. Furthermore, people who embrace the second conception often reject the first as being a correct conception of possible worlds. Most people who take there to be possible worlds that conform to the second conception deny that any of the merely possible worlds are large space-time regions with things like galaxies, tables, people or talking donkeys in them. Such people are called by David Lewis (1986c p 136) ersatzers about possible worlds, and the possible worlds they believe in are ersatz possible worlds, (since they believe in worlds according to the second but not the first conception: hence, to Lewis, they want to have second-rate or ersatz worlds).

However, it is possible to hold that both conceptions apply to possible worlds: Lewis is the most prominent example. His possible worlds are the same sort of thing as the world we live in (so the first conception is satisfied), and furthermore such worlds obey the correspondence condition, form a reasonably natural kind (‘maximally connected regions of spacetime and their contents’ or some extension thereof) and finally his worlds have a connection to modal truths specifiable independently of the correspondence condition: the specification is that the worlds and their contents make the modal truths true, and that this dependence of the modal on the worlds is what explains the holding of the correspondence condition. The correspondence condition is not part of the specification of truths at possible worlds: if anything, for Lewis it is part of the specification of what modal truths are. Others, such as some fictionalists and some eliminativists about possible worlds, may also hold that the first and second conceptions are equally applicable (i.e. both are only applicable to non-existent things, and then both are applicable to the same range of non-existent worlds). Thus the first and second

---

8 Eliminativism and fictionalism are not such good examples of positions which take the first and second
conceptions need not come apart, even though those who work explicitly with something like the second tend to reject the first as an accurate conception of possible worlds.

The heading of this section will have reminded some readers of Van Inwagen’s paper “Two Concepts of Possible worlds” (Van Inwagen 1986), and the two concepts of possible worlds he discusses (the Concrete conception and the Abstract conception) appear to be analogous to the two disambiguations I have mentioned. However, the two concepts of possible worlds Van Inwagen discusses are different from the two I have outlined in this section: I believe the distinction he wishes to draw should rather be identified with the concretist/abstractionist distinction I will discuss in Section Two (in fact I take it that I am following his usage in using the terms “concretist” and “abstractionist” in Section Two). The main difference between the distinction outlined in this section and the concretist/abstractionist distinction to be outlined is that the distinction I discuss in this section is a conceptual one rather than a substantial metaphysical one: philosophers can (and do) have things to say about possible worlds in both of my senses of “possible worlds”, regardless of whether they are abstractionists or concretists. Furthermore, it is plausible that people using “possible worlds” in the different senses I distinguish are just talking past each other, whereas it seems to me that they are more naturally interpreted as disagreeing about the same entities on Van Inwagen’s view. This makes it seem that I am discussing an ambiguity, whereas Van Inwagen is marking off sides in a metaphysical dispute. This is not to say that I disagree with Van Inwagen — the dispute between concretists and abstractionists is an important one. I think it would be a mistake to confuse that distinction with the one being drawn in this section, however. As for what Van Inwagen would say about the ambiguity I have diagnosed: at pp 192-193 of Inwagen’s article, he claims that the phrase “possible world” means something along roughly the same lines as the second of the two disambiguations I provide: it is a “functional concept” specified in terms of being entities suitable for various theoretical roles. I disagree when this is taken as a claim that this is the meaning of the term, though not if it is merely a claim that this is a meaning of the term. He appears to mean the former, though as he does not diagnose or discuss the ambiguity I have discussed it is unclear what he would say were it to be suggested to him that there is this distinction as well as his Concretist/Abstractionist distinction.

1.2 Two Problems for the conceptions of possible worlds

conceptions to come to the same thing, however, because of various problems in non-trivially comparing the ranges of concepts which do not apply to anything that exists. Lewis’s theory will suffice to make the point, however.
There are two problems with the conceptions of possible worlds given above that I wish to discuss at this point, since they are problems involving the conceptual analysis just provided. The first is more easily dealt with. It is widely thought that possible worlds provide the metaphysical grounds for modal truths of various sorts: that not only is there a biconditional between “necessarily Ø” and “at all worlds, Ø”, but also that when it is necessarily the case that Ø, it is so because it is the case that at all possible worlds, Ø. Is this part of the concept of what it is to be a possible world? It seems clear that it is not, in the first sense of the expression “possible world” — how other concrete cosmoi might or might not be need not explain the truth of modal claims. (David Lewis thinks that it does, but even he, I suspect, does not think that this is a straightforward analytic connection, but rather a substantial thesis to be argued for. His arguing for it as if it were a substantial thesis would be hard to account for otherwise).

In the second sense of “possible world”, it is not so clear that this is not part of the conception, but here too I do not think that it is an essential part of the conception. Even leaving aside the (doxastic) possibility that possible worlds and modal truths could be seen on a metaphysical par, with neither reducing to the other, it has always seemed clear to me that it is a substantial (rather than stipulatory) question whether possible worlds were the basis for modality, or whether it might be that possible worlds, and possible worlds discourse, were susceptible to reduction to modal discourse rather than vice versa. Furthermore, there is a lingering suspicion among many theorists that there are undischarged modal primitives (such as compossibility, co-satisfyability, entailment characterised as necessary truth preservation, or suchlike) lurking within many standard accounts of possible worlds. These are often thought to be unattractive parts of the theories which they infest, but I have never seen it claimed that this meant that the theories could not really be talking about possible worlds. However one would expect this claim to have been made if it were analytic that possible worlds, if there be such things, would be the reductive base of modal truth. Finally, “possible worlds” would be useful for many of the same tasks (e.g. formal semantics for natural languages, modelling of counterfactual talk and inferences, probability theory) even if they were not essentially the metaphysical ground of modality, and it is these roles in philosophical theory which many people feel are in practice the most important features of the things bearing the title “possible worlds”. A final consideration that argues in favour of the claim that it is not essential to the conception of possible world that such things be the ground of modal claims is that there are various theorists who both believe in possible worlds, of various descriptions, and yet do not think that modality reduces to them. Fine in (Prior and Fine 1977) is one example. These considerations together should show that possible worlds being the metaphysical ground of modality is not an essential feature of at least one perfectly acceptable conception of them, and I will not be taking it as essential to the
concept in this work. (Especially since I will not even be taking it that it is true that possible worlds are the metaphysical foundation of modality).

So much for the first issue. The second problem partially arises out of the answer which I have provided for the first. Recall that, on the second conception of possible worlds, it is plausibly a requirement that possible worlds have some salient connection with modality independent of their satisfying the biconditionals about necessity and truth at all worlds, and possibility and truth at some world, respectively. The second problem I wish to address is the problem that, in theories where there is a plausible candidate to be possible worlds in the second sense, there will usually be more than one such candidate which meets the biconditionals, is “natural” and non-gerrymandered, and yet has a salient connection with modality. When this is the case, is there any guide to help us choose between them, or are they all equally good (and perhaps therefore not good enough) candidates? I suspect that there is. In order to present what I take to be the principle at work, let me present a specific example, both to provide evidence for the claim that where there is one deserver (or collection of deservers, or collection-of-deservers-plus-relation) there are many, and to help to illustrate my suggestion.

According to David Lewis’s account, there are many concrete connected spatio-temporal (or somewhat analogously spatio-temporal) objects which exist — there are lots more objects of the same sort of thing that our cosmos is. These are possible worlds in the first sense of that term, and are, according to Lewis, also the possible worlds in the second sense of that expression. And, sure enough, they meet the three criteria I laid out before. But there are other candidates in Lewis’s ontology. For Lewis, properties are sets (he identifies properties and their extensions). Corresponding to each concrete world of Lewis’s there is the set that has that world as its only member. These singleton sets, taken together, (and where the “true according to” relation is taken to be something along the lines of “true in the member of...”) are appropriately related to necessity and possibility (at least according to Lewis, since their relations will be isomorphic to the relations to the modal truths had by their members when a suitable relation is specified), and are not hopelessly gerrymandered. Furthermore, they have a salient connection to modality. For possibility, for example, is intuitively thought to have to do with “ways a world could be”9. Lewis wishes to identify such ways with the concrete worlds, but it is perhaps more plausible to see ways as being properties — the way this cosmos is might be more plausibly identified with the property of being the way this cosmos is than with the cosmos itself. The singletons of Lewis’s worlds have a good claim to be the way those (concrete) worlds are: and since there is a conceptual link between the possibilities

---

9 As has been pointed out by Stalnaker (1976) and Forrest (1986). Forrest claims that this conception can be found in Husserl and Leibniz too.
and the ways worlds could be, there is a salient connection between the modal truths and these singletons.

There are other examples in Lewis’s system too. One of the candidates to be treated as propositions, for Lewis, are sets of concrete worlds — each proposition is the set of worlds at which it is true (so, for example, the proposition that talking donkeys exist is the set of all worlds containing talking donkeys). One common conception of possible worlds (in the second sense) is that possible worlds are maximally consistent sets of propositions. These have a fairly salient connection with modality — for at least insofar as modal operators are thought to be applicable to propositions (so that, e.g. it is propositions that are necessary or not) it is plausible to think that worlds might have a close connection to propositions — and according to the maximal-consistent-sets-of-propositions conception, a proposition is true at a world just if it is one of the members of that world. Another reason to think that worlds are propositional is that things are conventionally said to be true or false according to them. Representations are the only things typically that things are true or false according to, so a common line of thought is that possible worlds are representational. And what better to represent than propositions, or indeed sets of them (where a set of propositions can be thought of as derivatively representing the things represented by its contents)? So certain sets of sets of Lewis’s concrete worlds also have a claim to a salient connection with modality, as well as of course satisfying the other two requirements. (Since a proposition is possible iff it is a member of some maximally consistent set of propositions, and a proposition is necessary iff there is no maximally consistent set of propositions which contains its negation, the first condition is satisfied, and maximally consistent sets of propositions are not the sort of thing which are hopelessly gerrymandered or unnatural).

Thus there are at least three sorts of things in Lewis’s ontology which meet the three criteria for something to be a “possible world” in the second sense in which I outlined, yet intuitively it is the big concrete regions of spacetime with their contents which are the best candidates, rather than their singletons, or more complicated set-theoretic structures containing them as ur-elements. Perhaps there is not too much significant to this intuition, or perhaps it is a result of the fact that only the concrete candidates are “possible worlds” in both senses of the term. But if something of the intuition of a more general nature can be articulated, it may be useful as a further articulation of at least the second conception of “possible worlds”.

One suggestion is that set-theoretic constructions out of the concrete objects might be thought to be more derivative entities, or have a more derivative connection with modality. This is a hard intuition to justify, especially if one does not want to accept different levels of being for different sorts of objects. Standard modal accounts of derivativeness are not of much help either (in this case, as in many others, counterfactuals like “had the one sort not existed, the other sort would not have existed either” either do
not make much sense, or are trivially true because of the necessary falsehood of the antecedent). But if it is intuitive that sets of objects which have certain relational features which somehow involve the fact that they are sets of those objects depend on their members in having those relational features, then there is a sense in which the rival candidates to Lewis’s preferred candidate do have their connections to modality derivatively. The ways (concrete) worlds are, for example, are only those singletons whose members are the entities in Lewis’s ontology which are worlds in the first sense — and intuitively the singletons are ways worlds are just because their members are concrete worlds. And a set is not a proposition of the appropriate sort in Lewis’s system unless all of its members are concrete worlds\(^\text{10}\) — and so for a set of such sets to be a set of propositions intuitively depends on the concrete worlds. Unfortunately the situation is somewhat symmetrical — for instance, the only individuals which are concrete worlds are those which have one of the singletons which are the ways that worlds are, but there is not the same intuition that the concrete worlds depend on their singletons for being world-candidates in the same way as the situation seems to be vice versa.

Perhaps this notion of derivativeness, despite its intuitiveness, has nothing answering to it in Lewis’s theory, and in a sense this does not matter. What I wish to suggest as a further elucidation of the second conception of possible worlds is that, where there is more than one collection of candidates which meet the three conditions I previously suggested, then if one candidate’s status is derivative upon another’s, then the less derivative of the two has a better claim to the title of “possible world” (they are the better deserver of the two, to use the Ramsey-Lewis-Carnap jargon). It is enough to point out that the candidates in Lewis’s system which seem to be the least derivative seem to have the best claim on the title “possible worlds” in order to provide evidence for this further elucidation. Whether or not in fact they would be less derivative, or indeed whether they would in fact be the best deservers in Lewis’s system, is another matter.

Other cases besides those available in Lewis’s framework are even more clear. Take the following case: suppose that there is a theory of propositions under which the propositions are closed under negation (e.g. the negation of any proposition is itself a proposition), there is at least one necessary proposition, and that possible worlds are taken to be maximal consistent sets of propositions: all of which are common positions in the literature. Various permutations will be able to be performed on these sets of propositions so that the image of the maximal consistent sets, together with the appropriately modified relation, will satisfy the three criteria. Take for example the sets’ image under negation: so that each maximal consistent set will be mapped to the set containing all and only the negations of each proposition in the maximal consistent set.

\(^{10}\) Lewis’s preferred candidates to be propositions are sets of possibilia rather than sets of entire worlds — see Lewis 1979. A similar point can be made about sets of parts of the concrete worlds, of course —
These (in one sense) maximally inconsistent sets of propositions are plausibly saliently connected with modality — something is necessary if it is a member of none of these sets, and something is possible if its negation appears in at least one. These sets are plausibly thought of as the very opposite of possible, and with that construal it is plausible that the necessary truths could be found nowhere in them, and any negation of a possible truth will be found in them somewhere. There is at least some salient connection made out with modality through their satisfaction of the notion of “opposite of the possible”. However, if we were forced to choose between these sets, and the relation of “true in” which a proposition has to a world iff its negation is a member of it, and the maximal consistent sets of propositions with their “true in” relation of membership to be the possible worlds/ “true at” pair, it seems obvious (to me at least) that the maximal consistent sets and their corresponding relation deserve the name better. I hope it seems to the reader that the other candidate I defined appear less basically connected with modality, and that this is not merely an artifact of the way I introduced them — it is possible to introduce them without mentioning maximal consistent sets of propositions, just as it is possible to introduce maximal consistent sets of propositions without mentioning their “negated” cousins.

If one were unconvinced by my appeal to the admittedly presently slightly mysterious notion of “derivativeness”, one could attempt to discover some other intuitive criterion to function to select between competing candidates, or hope that one of the rival candidate (objects, relation) pair will score better on the three criteria of satisfying the correspondence condition, naturalness, and having a salient connection with modality, and so (perhaps narrowly) to do better than its rivals and be the best deserver for the title. I wish such a rival luck, but cannot believe that their chances are very good. Or one could think that there is no sufficiently principled way to choose between rival candidates: when one’s ontology serves up more than one competitor, it is either an arbitrary choice which competitor gains the title of “possible world”, or some structuralist account could be invoked (where we talk generally about all competitors, and something is true of the possible worlds just in case it is true no matter which competitor we select, false if true for no competitor, and there are various ways to treat the case of a claim being true for some-but-not-all of the competing collection of objects/ “true at” pair). This structuralism would differ from the structuralism considered earlier in this chapter, in that the additional constraints that the competitors be not too gerrymandered, and that they all have a salient connection with modality have been retained. I suppose that the first condition concerning the correlation between modal truths and what is true at possible worlds could be weakened slightly in some circumstances. One such circumstance would be if it turned out that there were several candidates which satisfied the naturalness and salience conditions, but none of which quite matched the correlation-with-modality condition, such sets seem to depend on the concrete cosmoi too in the same way that sets of entire “worlds” would.
especially if some principled story could be told about the cases where the candidates failed this requirement).

Finally, I suppose, one could think that if it could be shown that there was more than one candidate which met the required conditions, that would just show that nothing was a possible world (e.g. if one thought it was central to the concept that there only be one such collection of objects suitable for the role). This sort of eliminativism seems to me to be throwing the baby out with the bath-water, and it seems paradoxical to suppose that an embarrassment of riches should drive us into poverty in this manner. I suspect this eliminativism would not have too great an impact on much of what I had to say however — for it will provide a notational variant on a possible worlds account, according to which there are no possible worlds, but there are various objects/relation pairs which can be used for much of the work which possible worlds are desired — it is just that they do not deserve the name. (We might choose to call them, or one of the pairs among them selected arbitrarily, possible worlds* instead11).

So much for conceptual truths about possible worlds. Perhaps a more interesting question about possible worlds is what the correct substantive account of possible worlds is. There is, of course, much disagreement about what the correct account of possible worlds is, and while it is not my intention in this work to attempt to settle this question, it is important that I outline the alternatives. I will be making reference throughout this thesis to different varieties of accounts of possible worlds, and much of the time I will be concerned to address issues that only arise in the context of some accounts, but not of all. It is worthwhile then to provide a sketch of the philosophical terrain regarding the nature of possible worlds — and while this is certainly not the first such sketch, it will serve to introduce some terms for the various positions, and terms which I prefer on taxonomic grounds to the others I have found in the literature. In the section that follows, I will provide this taxonomy of theories of “possible worlds” in the second sense which I distinguished above. Indeed, for the rest of the thesis, unless context otherwise indicates, by “possible worlds” I mean possible worlds in the second sense of the expression, that is, by “possible worlds” I will mean those objects which satisfy the three conditions I outlined, and which are thereby fit to play the various theoretical roles expected of “possible worlds”.

11 Then, presumably, we could set up a convention which left the asterisk implicit in most of our usages. This would give the eliminativists the ability to translate their opponents homophonically when they want to produce more of the theory of possible worlds*. At this point, I have trouble envisaging such eliminativists’ quarrels with the believers in possible worlds being particularly heated.
2. The Competing Theories of Possible Worlds

There are four important distinctions between theories of possible worlds which could be incorporated into a useful taxonomy — for my purposes, I will concentrate on two and say less about the other two. The two distinctions important for my purposes are the distinctions between i) theories according to which possible worlds exist versus those according to which possible worlds do not, and ii) the distinction between theories which take possible worlds to be “concrete” and those which take them to be “abstract”. Before I discuss these, let me discuss the two distinctions I will not be taking into account to the same extent in this chapter. These are a) the distinction between those theories which assign realist truth conditions to statements about possible worlds and those which do not; and b) those theories which claim that modal statements are properly to be analysed in terms of statements about possible worlds, versus those that do not.

It is hard to spell out what realism is, or what its difference from anti-realism is. A (very) rough guide to one of the distinctions these terms have been used to mark out is that a discourse with realist truth conditions is such that the truth values of the claims in the discourse are mind independent — true or false independently of how we conceive of the subject matter, or facts about the workings of our minds or our epistemological apparatus, whereas an anti-realist account of truth conditions either requires some such mind-dependence, or alternatively denies that statements in the discourse are truth-apt at all — thus emotivism in ethics is anti-realist, not because it insists that the truth or falsity of moral statements is mind-dependent, but because it denies that such statements can be true or false. When applied to objects, typically realism claims that the objects’ existence or nature is not mind-dependent, whereas the anti-realist would either claim that it is, or claim that apparent talk about such putative objects is not talk about objects at all, but fulfils some other role (e.g. the emotivist’s claims about such putative objects as rights or duties, or the properties of goodness or evilness). Given this account of realism or anti-realism about truth conditions, one can be both a realist about a discourse (or domain of objects) and yet be an eliminativist — I am both an eliminativist about phlogiston and a realist about the truth-conditions of phlogiston theory, for example — I believe that whether or not there is any phlogiston, and if so what its properties are, is not a mind dependent matter, and yet I think there is no such thing as phlogiston. I mention this because it is sometimes thought that eliminativism about a subject matter or domain of putative objects is one way to be an anti-realist: but this broader notion of anti-realism is not the sense of anti-realism I am employing. In the sense in which I am using the expressions “realist” and “anti-realist” about truth-conditions or objects, this distinction crosscuts the distinction between theories which accept and those which reject the existence of possible worlds.
The questions of the pros and cons of realism and anti-realism in various areas are matters for interesting and important philosophical debate, but they are among the many issues which will not be examined in this thesis. I am temperamentally a realist about the truth conditions of statements about possible worlds (and indeed about most things), but much of what I say is, or can be, orthogonal to this debate, depending on what varieties of anti-realism are at issue. (Two important cases where it is not quite orthogonal — in my chapter on modal fictionalism I object to strong modal fictionalism for prima facie being committed to a fairly crude and implausible anti-realist account of possible worlds, and in my final three chapters I make free use of notions of infinity which some anti-realists about mathematics would take to be illegitimate.)

The other division in theories of possible worlds that I will not be discussing in detail in this chapter is the division between those who take possible worlds to provide the materials for an analysis of modality, and those who do not. I have indicated my leanings on this issue, but I will not be arguing directly for my preference in this regard in this work.

Now that I have mentioned some of the important distinctions between possible worlds I am not going to discuss in this section, let me move on to my discussion of theories of possible worlds in terms of those distinctions I do find important. I will begin by discussing theories of possible worlds which hold that possible worlds exist.

2.1 Theories According to which Possible Worlds Exist

Concretism

By far the most common position of those who have discussed possible worlds in the literature is to maintain that they exist (though whether this is because this is the prevalent view among philosophers, or is simply an artifact of the fact that people who do not think that possible worlds exist tend to not bother writing about them, is an open question). Belief in the existence of possible worlds allows for a pretty straightforward use of talk about possible worlds — in situations where it is useful to invoke possible worlds, one can simply mean what one says. Let me mention some of these views.

Perhaps the single best-known theory of possible worlds developed in the second half of this century is that of David Lewis's, according to which possible worlds are "possible worlds" in the first sense — some contain rivers and humans and galaxies and electrons just like the actual world. (Others, are, of course, more bizarre — as bizarre as it is possible to be, not surprisingly). Lewis's theory (most thoroughly developed and defended in his (Lewis 1986c), is the paradigm example of a theory according to which possible worlds are concrete. It is somewhat difficult to say what it is for an object to be concrete (as Lewis claims in (Lewis 1986c, pp 81-86), and which Van Inwagen, when
introducing his Concretist/Abstractionist terminology in (Van Inwagen 1986, p 188) also admits. It is especially difficult because the use of “concrete” does not even necessarily seem to be in accord with more usual uses — ghosts and Cartesian egos, to say nothing of things like romances and cricket matches, are all “concrete” according to this usage. (This partly depends on the rest of one’s metaphysical picture — at one stage at least Lewis took events, including presumably such things as romances and cricket matches, to be sets of various sorts (Lewis 1986b, p 244), and sets count as non-concrete for Lewis). Despite the fact that the divide between “concrete” and “abstract” is a difficult one to draw, it seems an easy enough distinction to grasp intuitively.

There are other accounts of possible worlds which take them to be concrete which differ markedly from Lewis’s in many ways. Lewis insists that no possible worlds overlap (or at least that they have no particulars in common — if there are immanent universals, they can appear in more than one world). A common variety of theory of possible worlds is that there are many concrete possibilities, but that they all overlap: these have sometimes been known as “branching” theory of possible worlds. These theories often seem to be aimed only at capturing some more restricted modality, such as some sort of temporal or perhaps nomic modality, rather than the more general “broadly logical” or metaphysical or logical modalities which most accounts of possible worlds also offer. The most notorious “branching” theory of possible worlds is the “many worlds” interpretation of quantum mechanics, at least that variety which is both intended to be realist and to be taken at more-or-less face value. This includes the classic original statement of the “many worlds” or “many universes” or “relative state” interpretation of quantum mechanics (Everett 1957). On p 147 of the paper in the pagination of Everett et. al. 1973 makes clear that he thinks that the other branches (worlds) besides the one we are in exist equally. In the note to p 147-148 he talks of all the branches being “actual” or “real” (apparently using the terms interchangeably) — he is not using “actual” in its technical sense. Typically each world or “world-line” overlaps with each one of the others, although one could conceivably produce branching theories with weaker unity, e.g. theories where each world-line is connected to each other only by an ancestral of the overlap relation, or even theories where there are distinct systems of overlapping world-lines. Some such systems, if they were to cover the gamut of possibilities rather than restricting themselves to more qualified modalities, will be more-or-less equivalent to Lewis’s system, with the main disagreement then centring on which world is actual. Lewis allows for worlds which branch (see Lewis 1986c p 209) — it is just that he does not identify the world-lines in these branching systems with worlds, nor does he think the actual world contains such branches.

12 Or we-as-we-take-ourselves-to-be, or somesuch. Everett talks as if he believes in literal identity between individuals in different “branches”, though his essay is not a metaphysical one and it is unclear
Abstractionism

The other category into which accounts asserting the existence of possible worlds fall is known as Abstractionism (following Van Inwagen 1986). This is not the only name by which it is known — Lewis applied the name “ersatzism” to such views in Lewis 1986c, with the insinuation that these “worlds” were inferior substitutes for his concrete candidates. To call such views “ersatz” is not entirely complimentary, but in this age of re-appropriating originally derogatory labels this, I suspect, would not be a barrier to the terminology being taken up. The terminology is, however, slightly misleading in one respect. It gives the impression that ersatzers are interested in producing an account which aims at an account of possible worlds in the first sense which I defined, and were driven to provide substitutes through not having such possible worlds available. But if all along they were only aiming for possible worlds in the second sense, their candidates are prima facie just as good as Lewis’s, and not in any way substitutes. It is because of this misleading impression, as much as for any other reason, that the terminology of “ersatz” worlds is less than desirable. “Abstractionism” as a name also appears to say more about what these views have in common — no matter what their differences, they take worlds to be abstract entities. I am unsure whether this informativeness is helpful, however, since it is unclear that “abstract” means anything more in this context than “not concrete” (whatever “concrete” might mean). As with “concrete”, however, “abstract” is a word which people seem to have little difficulty using in practice, even if explicit enlightening definitions were to be elusive.

Abstractionist accounts of possible worlds can be further sub-divided, into representational and non-representational accounts. According to representational accounts, possible worlds fulfil the possible-world role in virtue of being representational entities, and typically the biconditional connecting modality and possible worlds is something to the effect that something is possible iff some world represents that things are that way, either directly or indirectly (and necessary iff all worlds represent that things are that way, again directly or indirectly). Not all representational candidates for possible worlds are directly representational — I am classifying an account according to which possible worlds are sets of propositions as a representational account, even though it is not the sets that are the representational objects, but the propositions which are the members of those sets. Truth at such worlds is defined in terms of representation, however (a proposition is true at a world iff it is a member, or alternatively if it is entailed by the conjunction of the members, of such possible worlds). Accounts according to which the “true at” relation is defined in terms of representation, but in which the worlds are not themselves conceived of as representational entities in the usual sense, are accounts for which I will be using the description “indirectly representational”, and for what he would have said on issues such as trans-world-identity had he been properly aware of them.
which I will be saying that the possible worlds indirectly represent the claims which are true at them. Some standard examples of representational accounts include Carnap’s, one of the earliest accounts of possible worlds (called by him “state descriptions”) (Carnap 1956 p 9), Adam’s account, according to which worlds are sets of proposition (Adams 1974) and the Prior/Fine account of worlds as “world-propositions” (developed and defended in Prior and Fine, 1977).

Non-representational accounts of possible worlds, on the other hand, do not define the “true at” relation in terms of representation. They might claim (like Forrest 1986) that worlds are by-and-large uninstantiated universals which are comprehensive enough so that at most one could be instantiated at a time, and claim that \( p \) is possibly true iff for some such structural universal, it’s instantiation would entail \( p \). Or they might claim, as Plantinga does (or at least Van Inwagen claims that he does in (Van Inwagen 1986)) that possible worlds are certain existing but unobtaining maximal “states of affairs”, such that a proposition is true at a world iff were that world to obtain, the proposition would be true. Or such an account might be an extremely sparse account which says no more than that there is a world which a proposition \( p \) is true at iff possibly \( p \), and that a proposition \( p \) is true at all worlds iff necessarily \( p \). The distinction between representational and non-representational abstractionism is a distinction is worth making in its own right (it is not made clearly either in (Lewis 1986c) or in (Van Inwagen 1986) nor, indeed, in the literature at all so far as I know) and will become relevant in chapter 5, where the problem that alien universals raise primarily faces representational abstractionism but leaves non-representational abstractionism unaffected.

---

13 It is appropriate here to make a comment about Lewis’s terminology in his discussion of Ersatz accounts, so as to avoid confusion. In his general characterisation of ersatz accounts, he uses the idiom of a world “representing” something to be the case, so that every ersatzer believes in possible worlds being (at least) correlated to modal truths through “representation”. It is unclear from Lewis’s discussion whether he means to talk about representation in any full-blooded sense, in which case he has misclassified or ignored some abstractionist positions, or whether he rather is using words like “represents” or “representation” as place-filling vocabulary, his “represents that” relation holding between worlds and propositions being the converse of my place-filling use of talk about the “true at” relation. I suspect the latter, especially given his discussion on pp 174-191, where he explicitly mentions that some ersatzers use this talk of properties or states of affairs or whatever, with talk of “obtaining” or “realising” or suchlike, apparently taking these to merely be different ways of being a magic ersatzer, with these as rival candidates to be “representation”. Unless, for example, Lewis thinks in general that individuals represent the properties they have, and that this is what it is to have a property (which I am certain he does not), then presumably he must not be making very much of the word “representation” in his discussion.

14 Forrest’s account in his (Forrest 1986) is more detailed than this, and provides an account of the “truth at” relation (which Forrest in his account calls the “true under” relation) which does not rely on such modal idiom. Nevertheless, my modal characterisation of his “true at” relation is at least coextensive with his, according to his account, and gives the sense of the connection between his worlds and the modal truth.
2.2 Theories According to Which Possible Worlds do not Exist

As well as the above theories which endorse the existence of possible worlds, there are also theories according to which possible worlds do not. Clearly, some of these views will reject talk of possible worlds altogether — they are simply artifacts of a false theory, and to talk about them is no more helpful than to talk about phlogiston when one is interested in discussing things getting hotter or colder. (I will call such views eliminativist approaches to possible worlds). The approaches to possible worlds that reject their existence and which are most interesting for my purposes, however, wish to retain possible worlds discourse as being useful or appropriate in some (perhaps limited) way. Let me briefly mention some of these approaches, concentrating on those for which it is appropriate (or at least arguably appropriate) to think that they have realist truth-conditions.

One tempting way to deny the existence of possible worlds and possibilia but to nevertheless claim to make positive and true claims about them is to think that they can have objective characters without full ontic status — that is, for example, that possible blue swans are both blue and are swans, but have only possible existence. Such views fall under the broad category of Meinongianism. Meinongians can either ascribe a degree of existence which is less than full-blooded existence to possible worlds (such as Meinong himself ascribed to merely possible objects) or can deny that these objects have any degree of existence at all, while still claiming that they have characteristics, that our talk of possible worlds is talk about them, that our claim that “there are possible worlds at which it is true that there are blue swans” is still literally true, and so on. (A Noneist like Sylvan (Routley 1980) is a good example of a defender of this view). Meinongians can either be concretists or abstractionists (Meinongians about propositions or universals are clearly in an analogous position to those who defend the existence of such objects in being able to construct entities to play the roles of possible worlds). It is perhaps a difficult question whether Meinongians should be counted as realists about truth-conditions — the position is very similar in terms of mind-independence and objectivity, but introducing degrees of existence, or allowing for positive truths about non-existent

---

15 For the general taxonomic structure which I will use to divide different possible theories according to which possible worlds do not exist, I am indebted to (Nolan and Oppy, in progress).

16 The qualification about “positive” statements is to deal with the problem that not all the sentences of a given discourse will be false if both a proposition and its negation are sentences of that discourse. The claim that there are no possible worlds, for example, is a claim framed in the vocabulary of possible worlds. Yet one does not have to defend the legitimacy of talk about possible worlds to accept that claim. Indeed, it is precisely the defender of possible worlds that will need to claim that that statement is, if not false, at least illegitimate in some other way. The “positive” claims about possible worlds are typically those that commit one by asserting them to there being possible worlds. The “negative” ones are typically the negations of such claims.
objects certainly varies from the norm of realism. I suspect not much hangs on this demarcation question, and it is perhaps a question best settled by stipulation.

A second way to deny the existence of possible worlds and yet attempt to preserve the truth of claims which apparently involve asserting the existence of possible worlds is to be some form of contextualist. A contextualist about possible-worlds talk will claim that the surface structure of claims apparently about entities like possible worlds is misleading, and under paraphrase, or analysis, or when the statement is put in canonical form, or whatever, it turns out that possible worlds claims are really equivalent to more ordinary modal claims. This strategy has some problems — after all, it is not clear that a paraphrase can be found for every claim in talk about possible worlds (a famous example is the claim that there are at least two worlds). Contextualists will probably agree with eliminativists on the metaphysical issues involved, but it is just that their different views of the language involved give them reason to assign different truth values to the claims at issue. In some respects, contextualists about possible worlds are like those nominalists about properties who nevertheless think that claims like “Justice is a virtue” or “Red is more similar to orange than it is to blue” are true, because such claims, when subjected to the appropriate analysis, or understood in the correct way, or whatever, are not really about properties with names like “justice” or “red”. Of course, the contextualist about possible worlds will face some problems similar to those that may face such nominalists: if the contextualist thinks there is some way of taking “There are possible worlds” so that, suitably interpreted, it comes out as true, then there will be a similar problem to a nominalist who thinks there is a sense in which “There is such a colour as red” or “Red is a property” are acceptable: they may be stuck with claiming both that “X exists” is true, and that X does not exist, for some substitution instance of X, which is apparently just contradictory. This is not an insurmountable problem, of course.

One can also attempt to salvage the usefulness of assertions about possible worlds without thinking that those assertions are literally true. One method of doing this is to be fictionalist about possible worlds — to claim that possible worlds do not literally exist, but do exist according to a certain story or theory, and that this, together with some rules for acceptable inferences to and from modal truths and truths about the content of the story, provides a useful tool for many of the purposes for which possible worlds are useful. Fictionalism, especially fictionalism which purports to provide a basis for modality, faces some difficulties, as I will argue in chapter 4 and have argued elsewhere (see, for example, Nolan and O’Leary Hawthorne 1996). Fictionalism can be formulated in either concretist or abstractionistic terms — a good example of concretist fictionalism is that proposed by Rosen 1990, where the possible worlds are, according to the fiction, large concrete spatio-temporal entities like the worlds of David Lewis. Abstractionist fictionalisms are also possible, where one is fictionalist about sets of propositions, or structural universals, or such things. The fictionalism of Armstrong 1989, which
proposes a fiction which claims that there are many fictions each describing a world (in the first sense) in all of its detail, can be seen as an abstractionist fictionalism, where the possible worlds (in my second sense) are taken to be the many fictions each describing a possible world (in my first sense).

Of course, instrumentalism about possible worlds is a similar strategy, perhaps most saliently distinguished from fictionalism only by claiming that talk about possible worlds is meaningless or truth-valueless or to receive the truth-value “gap” rather than to be treated as being true or false. Instrumentalism, like fictionalism, will have some “prophylactic operator” (in the terminology of Nolan and Oppy) which, when applied to the neither-true-nor-false statements about possible worlds, turn the acceptable ones into truths. Some account of acceptability, plus import and export rules, provides a system which is supposed to provide many of the benefits of possible worlds. Again, it is not obvious that instrumentalism has realist truth-conditions, though again this depends on exactly what having realist truth-conditions comes to (and partially, I suppose, on how the instrumentalism is explained. Sentences which are not truth-evaluable certainly do not have realist truth conditions, but perhaps sentences which take a third truth-value such as “gap” might). Instrumentalism might appear to be able to come in both concretist and abstractionist varieties, but since it is (at least usually) intended to be an empty calculus rather than be meaningfully about domains of objects and their natures, this appearance is probably illusory. As well as whatever difficulties instrumentalism faces qua being instrumentalist, it also is heir to many of the same problems as fictionalism, especially if it is intended as an analysis of modality. No examples of instrumentalists about possible worlds spring to mind, but I would be surprised if there were none, or at least if it is a position which will not be occupied sooner or later.

This taxonomy is only intended to be a sketch of the intellectual terrain in this area, rather than a detailed description of all of the positions currently held (let alone all of the positions which it might be feasible to hold). In addition to the distinctions that I have drawn in this chapter there are many sub-varieties of each type of position I have mentioned, though by and large the details differentiating views within these broader categories will not be important for my purposes in this thesis (and when they are, of course, I will detail them at the time). In the next chapter I will address another set of conceptual issues which are important in philosophy of modality and possible worlds: I will address some questions of how to evaluate rival modal theories in terms of their theoretical commitments, and some issues connected with the question of what it might be to provide a satisfactory explanation of possible worlds or modality. The next chapter will make a start at providing the resources for explaining the difference between the two views. The main challenge in setting out this distinction is to explain what it could be for one to be explanatorily prior or metaphysically more fundamental than the other, and the next chapter at least makes a start at addressing that question.
Chapter 2 — Towards Explaining the Connection Between Modality and Possible Worlds

It is clear that modal talk (i.e. talk involving concepts like necessity or possibility) and talk about possible worlds are closely linked\(^{17}\). Much (indeed probably all) that can be said in the language of necessity and possibility can be said in the language of possible worlds, and much (though exactly how much is a matter of dispute: see chapter 3) of what can be said in the language of possible worlds can be said in the language of necessity and possibility. Or at least this seems to be the case, though objections to one or other way of talking may make it turn out that this equivalence is only apparent. In any case, even if there is not quite equivalence, if these two domains are closely linked we would like an account of what the connection between them is. To answer this question, I will examine some of the candidate varieties of connection which are thought to be explanatory connections between theories or their domains, and how applicable they are to this case.

Then I will go on to discuss another important issue relevant to explaining the connection between modality and possible worlds, and indeed to explaining connections between theories and/or their domains in general. People sometimes claim that either modality or possible worlds should be taken as primitive, and the one not taken to be primitive should be explained in terms of the one which is (plus perhaps some other theoretical resources). Furthermore, it is often accepted that the way to evaluate rival theories in terms of their theoretical commitments is to examine which theoretical resources are taken to be primitive by each rival. But what is a theoretical primitive? What distinguishes a theory which takes modality to be primitive from one in which possible worlds are primitive? Furthermore, some accounts of modality or possible worlds seek to explain them without appeal to “modal primitives”. But what might modal primitives be? The second half of this chapter will consist of an exploration of issues connected with the notion of theoretical primitives, and propose some suggestions as to how we should go about assessing theories of modality for the costs incurred by those theories in virtue of the primitives they employ.

---

\(^{17}\) In this section (as well as in others) I will sometimes talk of “possible worlds talk” or “the language of possible worlds” or “modal operator talk”, or employ similar expressions. There are many different languages (or language fragments) of each sort, especially if one considers the many formal languages available. Nevertheless, I will be employing the definite article and otherwise talking of such languages as if there is only one which talks of possible worlds and one which employs modal operators. This is not because I have a view about which instance of each type is the ideal, or anything: it is merely that the distinctions between the different regimentations of talk about possible worlds or different treatments of modal operators are by and large not relevant to the issues I am discussing here, and talking as if there is only one variety of each is simpler and less distracting.
1. What Might the Connection Between Possible Worlds Talk and Modal Operator Talk Be?

It seems clear that there is at least a correlation between the two ways of talking — the correlation expressed by the biconditionals connecting what is possible with what is true at some world; and what is necessary with what is true at all worlds. The question of whether the truth or acceptability of one way of talking supervenes on the other is exceptionally trivial, and so in terms of explanatory value is more or less worthless. Supervenience of As on Bs is cashed out in terms either of “necessarily, no difference in As without difference in Bs” or in terms of “for all worlds, no difference in what is true about As without difference in what is true about Bs”\(^{18}\). And given that the modal truths are plausibly thought to be necessary, and the truths about the existence and nature of possible worlds are plausibly thought to be true at every possible world (in each case this is pretty much the intuition captured by the modal logic S5, or rather the modal logic U), statements about what the modal or the possible worlds might supervene on are almost useless — for since there can be no differences in the modal truths and the truths about possible worlds cannot vary from world to world, both of these truths supervene on each other, and indeed on anything whatever. The truth that “if snow is white and grass is green, then snow is white” supervenes on my coffee mug, for example, because there can be no change in whether or not “if snow is white and grass is green, then snow is white” without a change in my coffee cup — and this is just because there can be no such change with or without anything at all.

Some further account of the connection must be offered. One tempting account is to provide an analytic reduction from one to the other (or establish equivalence and claim that the two discourses are on a par: perhaps through evolving in parallel). If there are things that can be said in the one language which cannot be said in the other, analytic equivalence looks dubious. While I argue in chapter 3 that more can be said in the modal language than some critics have given it credit for, it still remains controversial whether or not a one-to-one correspondence between sentences in each language can be found which would be even a \textit{prima facie} candidate for the relation of analytic equivalence between them. But there are other problems facing such a project as well. Possible worlds language, when taken at face value, apparently commits one to a domain of objects (possible worlds and perhaps \textit{possibilia} too), whereas modal language is not so clearly committing. \textit{(De dicto} modal claims, for example, are formed by attaching certain

---

\(^{18}\) Of course, supervenience claims are often both more precise and more subtle than the ones I have provided (they will give an account of what I’m calling “a difference”, for a start). There is an entire literature about supervenience and how to formulate supervenience claims, but I am not concerned to resolve the issues involving supervenience here, but only to indicate the particular problem of employing supervenience claims in this area.
intensional operators to sentences: and it is unclear what sort of commitment in general such sentences provide). In fact, some have argued for the rejection of possible worlds talk in favour of use of language relying on modal operators precisely because they think that modal operator language does not commit one to any new objects in the way that talk about possible worlds apparently does. Of course, the fact that some people do not see a putative conceptual connection is not decisive evidence that there is no such connection. It is reasonable, however, to think that the connection between the two ways of talking is at least underdetermined by our a priori grasp of the concepts. It may well be worth attempting other explanations of the connection.

It may well be thought that whether or not the two systems come to the same thing is not merely an analytic matter — for instance, it might be a matter for total theory, for example, where what we wish to say about the connection will be determined by what produces the most overall simplicity, or coherence, or completeness of explanation, or what satisfies some other global virtue in the most superior fashion (or of course some sort of combination of these). This holistic approach to evaluation of specific theoretical claims is of course Quinean in spirit. Or, indeed, some other method may be relied on which amounts to more than classical conceptual analysis of the sort that would be needed to justify an analytic equivalence. If such a procedure was adopted, the connections between the truths of claims in the language of possible worlds and claims cast in the language of modal operators would plausibly be more than semantic (just as, to use the classic example, talk of heat and talk of mean molecular kinetic energy are closely related, but their relation is not one to be discovered merely through conceptual analysis, and the identity of heat and MMKE in some cases is not a result of linguistic fiat). What sort of connection might be claimed that is not a semantic one between the two ways of talking?

Supervenience need not be a semantic matter — but supervenience has already been ruled out as not being a terribly helpful notion in this context. Another common non-semantic connection between different things is also not likely to be helpful — nomic connections between facts about possible worlds and facts about modality are clearly not something that can be usefully invoked to explain their connection. Some form of metaphysical reduction might be more helpful. Metaphysical reduction does not require that there be any systematic semantic correlation between different discourses or areas of the same discourse (such as translation, or even claim-for-claim correlation), and is not necessarily a matter to be discovered a priori, but beyond that it is a little difficult to say what a metaphysical reduction of one discourse to another, or one domain of entities to another, might amount to. Like supervenience, this is an area in which a great deal of philosophical work has been done, and probably still needs to be done, so I will not attempt any systematic or comprehensive examination of reduction here. Some suggestions for reductionist strategies, however, will be mentioned.
One method of reduction which seems to be reasonably well-understood and familiar is to reduce one domain to another by means of establishing *identities*. If every hydrogen atom is identical to a system consisting of a proton (with sometimes some attached neutrons) with perhaps one (or sometimes even two) electrons appropriately associated with it, then hydrogen atoms (originally dealt with primarily in chemistry) can be reduced to entities described in the language of particle physics. Similarly, hydrogen molecules can be identified with entities also described by particle physics. Various substances important to the study of biology, especially in the very small, can be described as compounds or mixtures of compounds able to be described in chemical terms—a reduction is in principle possible from biochemistry to entities described in the language of physics. Reductions through identity have been successfully carried out between various domains of chemistry and various domains of physics.

As well as connections between domains of entities, metaphysical reductions can be hoped for between different fragments of discourse. An at least partial reduction of discourses can be achieved when there is a reduction of the domain of one discourse into another, since co-refering terms can be substituted into, for instance, a fragment of chemistry so that the relations and properties are ascribed to entities picked out in the vocabulary of physics. But more than this can be achieved—as well as an identification of entities, accounts of the properties and relations ascribed in one fragment of discourse can be framed in the other. For example, “covalent bonding” between two entities can be cashed out in the language of physics—specifically, in terms of electromagnetic forces acting on electrons in the electron-shells of the relata. Here, too, I think it is tempting to seek reductions through seeking identities. The identities to be sought here, I suggest, are identities of *facts* or *states of affairs* which are differently described in the different discourse fragments, or (perhaps alternatively) identities of *truth conditions* of the sentences of each, or identities of *truth-makers*, or *a posteriori* identity of *propositions* picked out by the sentences of each, or some such. Nominalists, of course, will not have a bar of such identities—but I am tempted to think this is more of a crippling disadvantage of nominalism rather than a problem with this sort of reduction-through-identity project. Evidence that certain facts described in modal idiom were identical to facts described in the language of possible worlds would not be trivial in the way that a relation of supervenience would be, and general principles about which facts were which, plus a further metaphysical story (of the nature of possible worlds, or of the nature of facts more perspicuously stated in the language of necessity and possibility) would be the sort of account that could be a real improvement of our pre-theoretic understanding of these matters.

Some have thought that reduction, even metaphysical reduction, can proceed in a manner other than the establishment of identities. Perhaps something weaker than identity will suffice. One case where there seems to be reduction without identity (at least
in the case of domains) are the putative reductions of objects with proper parts to those proper parts. It does seem intuitive that in at least one sense mereological fusions are nothing over and above the (proper) parts that make them up — it is plausible that one needs no more than a thing’s proper parts to exist for that thing to exist as well. (This can of course be denied — it might be thought that the thing which is me would not exist at all if my atoms were scattered — but there is at least some prima facie plausibility to thinking that this thing would exist, even if it were scattered, and so not a human body, let alone a person). However, there cannot be identity between a thing and its proper parts (as opposed to the aggregate of its proper parts): one of the most obvious problems is that a given thing is singular — there is only one thing which is it — while its proper parts, if it has any, must be plural — they might be two, or seven, or continuum-many. The whole-part relation then, seems to be a relationship other than identity which can ground reductive claims. There may well be other such relations — let us give them the generic title relations of constitution. The relation of universal-to-particular, or universal-to-state-of-affairs is an example of a putative constitution relation which is plausibly other than mereological connection. Constitution relations offer more flexibility in making reduction claims, at the cost of being potentially more mysterious. They also potentially have another attractive feature: they need not be symmetric in the way that identity must be. Reduction is supposedly asymmetric, so those that want to justify discovery of identities as reduction should ideally have to say something more about why there is reduction one way rather than the other, given that identity is a paradigm of a necessarily symmetric relation. Advocates of constitution, however, may be able to satisfy the asymmetry of reduction simply through pointing out the asymmetry of their constitution relation — the fact that a whole is made up of its parts does not invite the suspicion that parts must be made up of their whole, for example.

The connection between possible worlds and modality, then, is one which I hope will be illuminatingly explained by means of a metaphysical reduction, whether through discovery of identities or through an explanatory constitution relation, though the mere fact that the other alternatives I have examined are not satisfactory does not establish this, since there could of course be other forms of elucidation which I have not discussed, or are even perhaps as yet unthought of. In any case, I am not going to attempt a reductive explanation of either possible worlds or modality in this thesis. I do wish to perform a preliminary explanation of another issue connected to the explanation of modality and possible worlds, however, an issue which in a sense is of a level higher of abstraction. The version of explanation of possible worlds or modality I have suggested is most appropriate is some form of metaphysical reduction. But whatever form the "explanation" of a phenomena such as modality or possible worlds takes, (if indeed there is such an explanation), if such explanation is asymmetric then there will be questions of what is explained in terms of what, and where the explanations end. Assessments of
rival explanations in terms of the theoretical resources invoked will still need to be carried out. To deal with such matters an understanding of theoretical primitives is needed, and yet explicit discussions of this topic are rare. In the second half of this chapter, then, I will be addressing the question of what theoretical primitives are in general, and more specifically what modal primitives might be.

2. The Primitives of a Theory

Theories usually have a function of explaining some of the phenomena or facts that they discuss. The truth of some claims will be explained by making other claims, which are (if, but not only if, the theory is successful) perspicuously connected to the explananda in such a way that genuine explanatory progress is made. Eventually, however, most (perhaps all) theories make commitments, or rely on claims, that are not themselves further explained in the theory (or at least not further explained in some sense or other — there may well be heuristics to guide the acquisition of the needed concepts, or relations of discussions to the world by means of appeals to perception, or ostension, or something similar). When a theory makes a commitment to something's being the case which is not susceptible within the confines of the theory to further (reductive) explanation, that theory has a primitive commitment in virtue of being committed to that thing's being the case.

Two things need to be noted to make the above paragraph plausible. The first is that I am employing a theory-relative notion of primitiveness of a commitment, rather than an absolute notion. For instance, the primitives of some branches of biology might be cells and various non-living but important chemical compounds, even though those very things might receive a further analysis into more basic explanans — for instance biochemistry might analyse the compounds and the constituents of cells in terms of chemical atoms and the relations between them, and so the compounds and the parts of the cells would not be primitives of that biochemical theory. There is an absolute use of the concept of primitiveness of theoretical commitments, but it can be understood as a special case of the theory-relative sort — absolute primitiveness can be thought of as the primitives of one's total theory, or the primitives of the "most basic" of a hierarchy of theories (which would normally be metaphysics and/or physics, I suppose). There is probably also a sense of "primitive" where it can be meaningfully asked about what is metaphysically or ontologically primitive in the sense of what the primitives in the world are, and not merely what the primitives of a theory are. This too can be made perfect sense of — it can be thought to be a question about what the primitives of the One True

---

19 It is clear, but perhaps worth emphasising, that in saying this notion is a 'theory-relative' notion of primitiveness, I am not saying that what the concept of primitiveness is, or which conception is the conception of primitiveness, is relative to what theory one holds.
Theory are, or what the most basic things or facts in the world are (in the sense of being the reductive base for all of the others).

The second thing that needs to be noted is that I used "explanation" in a rather specific sense in the previous paragraph — exactly what sense is difficult to say, but it does not include "explanations" such as listing the causes of something, or listing teleological ends of a system of which the things or the facts are (perhaps improper) parts. A theory does not avoid a commitment to a primitive commitment merely by asserting that the thing in virtue of which it has a primitive commitment is caused by something else. (A vitalist, for example, cannot avoid commitment to distinctively biological primitives merely by asserting that every living organism has a parent). The only sort of explanation which dispenses with primitive commitments, it seems to me, is reductionist explanation (whatever exactly that may be). This is not entirely uncontroversial, I suspect — some of the more puzzling claims in the literature of the supervenience of the mental on the physical seem to amount to the thought that a supervenience claim, even there is no prospect of providing a corresponding reductionist claim, is enough to dispel primitive commitment to the supervening entities or processes or states of affairs. Nevertheless, I feel confident enough to take a stand here about what the concept of "primitiveness" amounts to, at least the version of the concept which I am interested in, if there is more than one variety.

Those that think that the only explanation is causal explanation might now be having trouble understanding what kind of "explanation" I have in mind. If their view is merely a view about what they will use the word "explanation" for, but is not inconsistent with allowing that there are also elucidations or informative restatements or important theoretical connections which any satisfactory theory will invoke but which are not, by their usage, strictly speaking "explanation", then those people should not take me to be talking about what they talk about when they use the word "explanation". One could also think that there is nothing to explanation except causal explanation and not be worried by what I say on a sufficiently generous account of causal explanation, where there is much more to the causal explanation of something than merely providing a list of events or things and asserting that these are the causes (or some of the causes). Causal explanation might be thought to be rich enough to have a place for reduction claims, or supervenience claims, or even other things like mathematical truths — such theorists should take me to be talking about one sort of what they take to be causal explanations. Finally, there are the hard-core "explanation is nothing but listing some causes" types, according to which I am not talking of explanation at all. So much the worse for them — I cannot guess what they might think statements of reductions or superveniences could be doing in an explanatory theory, and I suspect that if pressed they would have a hard time being able to explain what a primitive was — provided they could even allow themselves to understand the concept.
I have claimed that a theory incurs primitives when it is committed to something’s being the case that has no further explanation (in the special sense indicated) within the theory: but I have so far said little about what primitives might be. Fortunately, there is a common distinction between two sorts of primitives which a theory may be committed to, a distinction which as far as I know is exhaustive: primitives divide into ontological primitives and ideological ones.

A theory’s ontological primitives are the commitments to the existence of objects it makes, where it does not also state that such objects can be reduced to other objects which the theory is also committed to. This account of ontological primitives may deliver slightly surprising results when theories less than total theories are examined — since economics is apparently committed to minds (through being committed to agents having beliefs, preferences, and so on) and yet naturally says nothing about whether these minds reduce to anything more basic, it may appear that economics is committed to minds as ontological primitives. However, it is not the case that believing sections of economic theory should be taken ipso facto to mean that someone thinks the mental is ontologically primitive: it merely means that if they are reductionists about the mental, their reductionist claims are not part of their economic theory — when we ask about the theorist, we need to consider their theory as a whole to determine what entities they take to be ontologically primitive. Another point at which this account of what it is for a theory to be committed to ontological primitives seems to be in tension with intuition is that a person can apparently believe that entities which are not further reduced in their theory are not ontological primitives. A physicist, for example, may believe that quarks are themselves reducible to more fundamental particles (or at least may not believe that they cannot), even though that theorist may believe that such further particles have not yet been discovered (or at least believe that it is an open scientific hypothesis that such further particles might be discovered). Yet according to the account offered, it appears that the physicist is committed to quarks as ontological primitives.

In the case where the physicist positively believes that quarks are reducible to other entities, I could save my account of ontological primitiveness by claiming that quarks are not the ontological primitives of the physicist’s theory, even by the lights of the analysis I offered. For the physicist is committed to further entities — admittedly entities of we-know-not-what kind — that is just what it is for the physicist to believe that there are entities which we know very little about but which are the entities to which quarks will be reduced to. A theory can be committed to entities without saying much about them, after all, and there need be no requirement that the ontological primitives of a theory are described in very much more detail than any other objects which the theory claims exist. The case in which the physicist is agnostic, however, can not be so dealt with. For in that case the physicist is not committed to objects to which the quarks reduce, even objects of we-know-not-what kind: it is just that the physicist is not
committed to the denial of the existence of such objects either. According to the account just given, the physicist's theory has quarks as ontological primitives, yet the intuitive thing to say in this case is that the theory leaves it open whether or not quarks are ontologically primitive. This seems like clear evidence against the account I have offered.

To see that it is not, we have to remember that the notion of "primitiveness" comes in a theory-relative variety and a more objective variety. In one sense of primitiveness, the ontological primitives (simpliciter) are those entities which really are the reductive base for everything else — in reality, rather than according to any specific theory. Alternatively, the ontological primitives in the absolute sense are those ontological primitives (in the theory-relative sense) of the One True Theory which provides (or would provide) a complete account of reality. When the physicist is agnostic about the ontological primitiveness of quarks, it is natural to see that physicist as not wishing to affirm or deny that the ontological primitives in the absolute or objective sense are quarks — and in this sense of what it is to be an ontological primitive, the account I offered can (and should) agree. The account I offered should be seen primarily as an account of the theory-relative notion of ontological primitiveness, and it is even consistent with this account to hold that the absolute or objective account of primitiveness is the one in more common use. (In fact I suspect that this is the case).

There is still the issue of whether we should say even in the theory-relative sense that the physicist's theory has quarks among its ontological primitives — and I am prepared to admit that even when the absolute notion of primitiveness has been distinguished, it still sounds a little strange to say that the physicist's theory has quarks among its ontological primitives. I am inclined to say that this is a slightly unusual application of our conception of the primitives of a theory, but that it is still appropriate. Not much hangs on this issue however, since even if it turned out that I was proposing an novel extension to our conception I would still stand by it, as giving a useful handle on what it is for some postulated objects to be the ontological primitives of a theory. We need this notion as well as the absolute one if we wish to compare theories easily with regard to what ontological primitives they are committed to, and even if it was the case (as it often is) that our conception of what this comes to is not quite in accord with an explicit and perhaps artificially precise definition which I have offered, this does not mean that the proposed definition is worthless. In such a case, the "definition", rather than a report on

20 Of course, what the ontological primitives in the absolute sense would be if there was more than one non-equivalent totally comprehensive and true theory is an interesting question. Perhaps we should say that there are no ontological primitives in such a case, or perhaps the ontological primitives are the intersection of the ontological primitives of each non-equivalent totally comprehensive and true theory. I mention this not because I think that it may well turn out that there is more than one non-equivalent totally comprehensive and true theories (I do not even think that it is logically possible), but only because some theorists believe that it may well be so. My conceptual analysis of what it is to be an ontological primitive need not rest on my beliefs about the prospect of equally true and completely
actual meaning or usage, would more usefully be seen as a proscriptive recommendation for use. The line between reportage and recommendation in this sort area is certainly vague — I hope I have managed to land on the former side of the line, but it is no tragedy if I have succeeded in only achieving the latter.

Before leaving the issue of whether my account of what it is to be an ontological primitive (both in the theory-relative sense, and in the absolute sense) is correct, it is worth mentioning a rival proposal, which has the advantage of agreeing more with our intuitions in the case of the agnostic physicist, but the disadvantage of being somewhat messier. This account agrees with the one just offered about what it is to be an ontological primitive in the absolute sense. However, it gives a different account of how to tell what ontological primitives any given theory is committed to (that is, it differs in its account of the theory-relative notion of ontological commitment). According to this proposal, what ontological primitives a theory is committed to is determined by what the ontological primitives (in the absolute sense) would be if the theory were true. After all, a common part of assessing theories is to evaluate them via examining what would be the case were the theory correct. This method does apparently deliver the result that we do not have to say that the agnostic physicist's theory has quarks as ontological primitives, since for all the physicist has said it has not been ruled out that the theory could be true and quarks not be ontologically primitive in the absolute sense. There are at least two things that need to be clarified — the first is that when we are determining what would be the case were the theories true, it would be almost certainly necessary to interpret the theories so that their designators were non-rigid.21 This is because otherwise the natural-kind terms in such theories would mean that the kinds discussed would have their actual nature no matter what possible situation was being examined, rather than the nature ascribed to them by the theory. For example, Aristotle held that water was homogeneous and infinitely divisible. Yet it is not even possible that Aristotle's theory could have been right, given that water is necessarily H₂O. If we wanted to say both that Aristotle could have been right, and that if Aristotle was right water would have been homogeneous and infinitely divisible, we ought in this context to take "water" as it occurs in Aristotle's theory as not being a rigidly designating natural-kind term, even if that is the usual semantics we should wish to apply to that word (or its Ancient Greek equivalent, in this case). Too many theories would turn out to be not usefully susceptible to this analysis due to their being necessarily false if the a posteriori necessities which are supposed to follow from the employment of natural-kind terms are taken into consideration.

The second point is to notice that there will very often not be a unique way that a theory could be true. The agnostic physicist's theory could be true in worlds where

---

21 Unfortunately the term "flaccid designator" never seemed to gain much philosophical currency.
quarks (or rather quark-like things, if we are interpreting the natural-kind terms non-rigidly) are the ontological primitives, and it could be true in some worlds where the quark-like things reduce to something more fundamental. Indeed, this may well be true in most cases where a theory could be true — they could become true in worlds of all different levels of ontological complexity. This is I think the main defect of the proposal — there will almost never be one unique set of ontological primitives which a given theory will be committed to. This is likely to mean that there will be no fact of the matter about what primitives most theories possess. This may (perhaps) have the virtue of according better with usage and linguistic intention, however I believe that even if this were so this rival proposal should be replaced with the one which I first presented. In any case, my usage in this thesis will conform to the usage outlined in my first proposed account of ontological simplicity.

Specifying what it is for a theory to possess ideological primitives, and how to assess what ideological primitives a theory has, is a more difficult and probably more controversial matter. Ideology, as it is used in this sense, is concerned with the predicates and operators which a theory employs. As a rough first pass, a theory incurs an ideological commitment when it is committed to a true claim which involves the application of a predicate or operator. A theory incurs a primitive ideological commitment through employing a predicate or operator which is not further defined in terms of more basic predicates, operators, or ontological commitments. One way this account could be made more precise is to spell out what predicates or operators are. I will not attempt this here, except to point out that I am taking "operator" to include sentential operators, sentential connectives (such as the conditional) and possibly even other functors. As the account stands, it is rather linguistic. Predicates and operators are linguistic entities, yet the interesting aspect of theoretical commitments tend not to centre on what linguistic conventions we have employed to state the theory, or the connections between such linguistic conventions. Perhaps it is better to say that the ideological commitments are not a matter of what predicates or operators are employed and how they are defined, but in terms of the assertions produced through the employment of the relevant predicates or operators made by the theory, or some such. Exactly what one wants to say about this will depend on one's theory of theories (whether they are most perspicuously seen as sentential entities, or propositional entities, or abstract models of some other sort, or whatever). I will not try to sort out the details about what exactly to say in this matter here — I merely wish to point out that the issue of ideological commitment is not primarily a linguistic matter, despite the talk of predicates and operators. (Or at the very least it is not clear that this issue is primarily a linguistic matter).

One feature of the first pass provided is that it talks as if it is useful to talk about some predicates and operators (or what is associated with those predicates and operators)
being more basic or less basic than others. It is a tricky matter to spell out exactly what this might come to, especially in the case of natural languages. Taking the case of predicates, it is plausible to say that one predicate is less basic than another (or a collection of others) if that predicate is explicitly defined in terms of those others. While this may often be true, it does not seem however to be either a necessary or sufficient condition for one predicate to be less basic than another (or some others taken together). Not necessary, for it seems that a predicate could be less basic than another even if no explicit definition had ever been offered for the former: for instance, suppose that the former could be definitionally captured by the latter, but as a matter of historical fact has never been done so — it is intuitive that in this sort of case the former might be less basic than the latter. Not sufficient, because it might be thought that which predicates or concepts are more basic is a less artificial matter than the issue of what is defined in terms of what. Systems of definitions can often be set up so that intuitively unified predicates or concepts are defined in terms of apparently more gerrymandered ones. If “.. is red or round” and “.. is round” are available predicates, then “.. is red” can be defined from them, provided we are allowed the use of logical connectives in definitions (something is red iff it is red or round and is not round). Yet even if a toy language was constructed so that words for “is red or round” and “is round” were introduced first, and the word/s for “is red” was defined from those words, it might still be thought that the predicates “is round” and “is red”, or their equivalents, were more basic than “is red or round”. There are other worries about the link between definition (or definability) and the basicness of one predicate rather than another: a minor one is that the account of basicness in terms of definability does not intuitively get it quite right when there are two unstructured predicates which mean the same thing. Each can be defined in terms of the other, but it would not do to say that each is more basic than the other. They seem equally basic — or rather, it seems that we ought to speak of their being only one piece of ideological commitment gained by using both of them, rather than two. How problematic these problems for tying basicness to definability are is probably a controversial matter, but even the controversy is enough to indicate that “basicness” is not automatically tied to what as a matter of fact is, or could be, defined in terms of what.22

22 This paragraph and the next will indicate that I am not going to follow Goodman’s account of primitive predicates (Goodman 1966 chapter 3). I have not mentioned his account of primitive terms in my discussion of ontology at all, since his discussion proceeds entirely in terms of predicates. Goodman’s account of what it is for a predicate to be primitive in a system is simply that the predicate is not further defined (or definable) in the system: thus it is both a semantic conception of primitiveness and, at least in the context of his overall constructivist approach, is rather anti-realist. The result of Quine 1954 that any system of primitive predicates can be represented with a system containing just one dyadic predicate, provided more ontology is allowed, also casts doubt on the plausibility of Goodman’s account. Recently Hazen 1996a has provided another result that any system can be represented employing only a dyadic predicate, and as this result does even better by Goodman’s lights and is even more unintuitive, the plausibility of Goodman’s account is very suspect even if one shares many of Goodman’s
In fact, the measure of ideological commitment may well not be a matter of the number of predicates or the connections between the meanings of predicates employed at all. One might think that if exactly the same situations are represented by two theories which in addition possess the same ontological commitments, the theories have the same ideological commitments even if there is no natural isomorphism of the predicates of the one theory and the predicates of the other — perhaps even if it is impossible to provide a definition of the terms of one in terms of the other (although slightly rougher connections such as "rough translation" or "explication" might hold between predicates of the one and predicates of the other). In the case of reduction in general there is a conception of metaphysical reduction, where the relation between the objects or truths to be reduced and the objects or truths to be reduced to is not necessarily one that can be discovered by noticing semantic links between the claims about the domain or theory to be reduced and the domain or theory to be reduced to. Equally, there is such a conception when it comes to ideology: the relation between a predicate or group of predicates and another may be such that one is more fundamental (in a metaphysical sense) than the other even if there are no systematic semantic links between them (such as links of definability).

In the case of predicates, the account of what sort of relation might hold between different predicates is couched in terms of an account of the relations between the properties or relations associated with each predicate (the property or relation "expressed" by each predicate, as it is sometimes put), and the account usually takes the relation of being-more-basic-than between predicates as being correlated with the relation of being-more-ontologically-primitive which holds between the properties associated with such predicates. This method, however, will hardly be attractive to nominalists. Furthermore, such an account of the relation of "basicness" between predicates will not be acceptable even to most realists about properties and relations as a completely general account of basicness of predicates. For one thing, there tend to be some predicates which even realists do not wish to account for in terms of properties and relations. (So called class-nominalists, for example, tend to take the set-membership "relation" to not be a relation in the ordinary sense — for a set of ordered pairs whose range included all of the sets leads to contradiction in many axiomatisations of set theories). Secondly, to explain the connections between predicates by means of the connections between the properties associated with the predicates is, it seems to me, to reduce ideological commitments to ontological ones. (After all, it seems that once one has said which property a predicate is associated with, and what connections that property has to objects (including modal and constructivist and extensionalist intuitions.

---

23 Lewis 1983 p 353 claims that it is impossible to do away with unanalysed predication (without "artificial tricks" which only dubiously dispense with unanalysed predication). I do not agree with him — but I think I agree with the weaker claim that in fact no-one in the literature has done so. This is to say that everyone has ideological commitments which do not get explained or reduced in terms of their
hyperintensional connections), you have said all that could be usefully said through employment of the predicate itself). While this might often be appropriate — and, as a realist about properties and relations, I have no axe to grind against it — it will not help as an account of what ideological commitments are primitive, since the primitive ideological commitments will not be explicable in terms of ontology, even if many other commitments are.

More needs to be said about the connections between different collections of predicates when one collection is more or less basic than the other. In particular, so called Ostrich Nominalists like Michael Devitt (see Devitt 1980) owe us an account of the connection between more basic and less basic predicates (between the "fundamental" predicates and the other predicates that are "constituted" from them, to use the vocabulary of Devitt 1980 p 436). Of course, those who are committed to ideology that is not further explained in terms of ontology may wish to deny that any of it is to be explained in terms of other ideology either — I take this to be taking all or at least most of the ideological commitments of the theory to be primitive — but this is not open to anyone who wishes to be thought a reductionist about anything whatever. Devitt, for example, as a card-carrying physicalist, should have no truck with such primitivists.

This is of course also true when it comes to other linguistic features such as operators which feature in the claims of theories. An account of what it would be for an operator to be more or less basic than another, or more or less basic than a predicate or collection of predicates, is in some ways even further from being developed than an account of such connections between predicates, where the application of those predicates is not to be explained by properties or relations. Let us for now just assume that there is some sort of explanatory connection, which, when we are discussing metaphysical primitiveness in any case, may not be always straightforwardly connected with semantic connections such as what is in fact defined in terms of what, or even what is definable in terms of what. Some ideological commitments are able to be explained away in terms of ontological commitments, it is often assumed: that is, the application of a predicate or an operator can be explained in terms of what sorts of things the theory claims exists. Some others, however, are not (at least in most theories) — and the primitive ideological commitments of a theory are to be measured by examining those predicates and operators whose application is not to be explained (in the relevant sense) in terms of the theory's ontological commitments (or at least not entirely in terms of them), and which are also not be explained in terms of the application of other predicates or operators.
This account of ontological primitives and ideological primitives has not been exhaustive, and nor was it intended to be: but enough has been said about these concepts for my purposes. It has been established, at any rate, what sorts of things need to be considered when one is considering what the primitives (in the metaphysical sense) of a given theory are. With this conceptual apparatus, we are now able to consider the issue which is more directly relevant in this context: what it is for a theory to have modal primitives. This question is undoubtably an important one, for one of the tests of the attractiveness of a theory is the sparseness of its postulation of primitive ontology and ideology. When it comes to theories of modality, if one theory can explain modality at less cost, it is prima facie to be preferred. However, determining what primitive commitments a theory of modality postulates, and which modal commitments of that theory are susceptible to further explanation, can sometimes be a troublesome matter.

The issue I am interested in here is how it might be determined what the theoretical cost in terms of primitives of any given explanation of modality. This means that the notion of "modal primitives" that I am interested in is different from the notion as it is sometimes used in the literature. Sometimes, a theory is taken to be committed to modal primitives iff it takes a necessity operator or possibility operator as primitive. This sense explains, for example, why it is only those people who do such a thing which are commonly described as "primitivists about modality". The sense in which I am using the conception is broader — for instance, a theory which analysed modal operators in terms of possible worlds conceived of as sui generis abstracta would also count as being committed to modal primitives — the modal primitives being the worlds (and perhaps the "true according to" relation postulated as well). I will be using the phrase "modal primitive" and similar phrases, both in this section and in the thesis in general, to capture this aspect of evaluation of modal theories, rather than the comparatively uninteresting question of whether the theories take one or other standard modal operator as primitive or not. Given my usage, it becomes an interesting question whether it is possible to have a theory of modality without any specifically modal primitives: intuitively, whether or not modality can be accounted for using only those resources already available from the explanations of the non-modal. On the alternate usage, of course, this is a trivial question, for it is very easy to produce a candidate account for modality in terms of the "non-modal", if all that is meant by "non-modal" is a primitive other than one of the standard modal operators.

I feel I also must resist the temptation to employ a conception of "modal primitives" which, while prima facie tempting, would trivialise the issues of what the modal primitives of a given theory are, and what the modal primitives of the best theory should be. This is to say that the modal primitives are all and only those primitives...
employed in a theory of modal truths (or of modal phenomena, or modal facts, or however one wants to generalise about things modal here). Until now I may even have given the impression that this is what was meant by the phrase — this impression was unavoidable to an extent since I felt it necessary to put the other parts of the picture in place before ruling out this interpretation of that phrase. In hindsight the reader should find that I have been at least studiedly ambiguous about whether this is what I meant or not. One reason why this reading of the phrase “modal primitive” is not useful is that it might turn out fairly straightforwardly that all primitives, without exception, would be modal primitives. Anything which is the case is also possibly the case, and actually the case. If it turned out that the resources needed to explain something’s being the case were required to explain its possibly being the case, or actually being the case, if this is taken to be a modal matter, then we would require any primitive invoked in the explanation of anything that is the case in order to explain all of the modal truths, and so by this definition every primitive employed in the theory would be a modal one.

Even if it did not turn out that all the primitives of a theory needed to be invoked to account for the truth of modal claims, the definition of modal primitives as those primitives employed to account for modality would still be unsuitable. This may be because some of the primitives employed to account for modality may not have anything distinctive to do with modality, and may equally be important in an account of non-modal phenomena. To insist that such primitives are modal primitives would be like claiming that physicalists about the mind are committed to mental primitives, or primitive mentality, on the grounds that they too are committed to primitives in the account they provide of the mental — it is just that these primitives are the primitives of physics as well. If it is not appropriate for dualists or idealist to reply with *tu quoque* to the charge of having postulated mental primitives (and I assume that it is not), then there seems at least the conceptual possibility that one might be a non-eliminativist about modality and yet not be committed to modal primitives. A better measure of commitment to distinctively modal primitives is, I suggest, to examine what primitives are added to a theory *primarily* to provide a basis for the truth of modal claims, or alternatively to consider as modal primitives only those commitments, whether ontological or ideological, which are postulated to explain the truth of modal claims, and which would not be postulated in the theory otherwise. This accords well with what we would want to say about, for instance, the case of mentality — it is only those people who postulate specific ontology or ideology which would not be postulated anyway in, for instance, a physical or biological description of an organism who we would be tempted to label a dualist about the mind.24

24 On the assumption that the physical or biological explanations are not Idealist, or panpsychist, or whatever, of course.
I claim that the most useful way of assessing the cost in terms of the commitments of a theory of modality is to examine what the modal primitives of that theory are, where that amounts to an examination of those primitives postulated by the theory which are not independently justified in the theory to provide an explanation of some non-modal phenomena. This criterion itself is somewhat imprecise, and determining when a theory’s postulates are for the purpose of accounting for modality may be a matter of judgement. For example, a theory which is supposed to postulate very little in the way of modal primitives, but which has a lot of apparently fifth-wheel primitives dragged in to explain some non-modal phenomenon but which turn out to be vital for the account of modality offered by a theory may perhaps be better seen as underhandedly avoiding postulating primitives as modal primitives which it ought: but whether any particular case is such a case as this is bound to be a matter of controversy. It might also be a matter of controversy exactly how to describe such a case: whether it is that the theory has avoided employing modal primitives, but would be better if it did, or whether rather the supposedly non-modal primitives are not really non-modal, because their use in the explanation of the relevant non-modal phenomena is not justifiable. These difficulties do not render the criterion useless, of course.

Another, more serious, difficulty with this criterion is that it assumes that a useful line can be drawn between the modal and the non-modal. This may not always be uncontroversially the case. Counterfactuals, for example, seem to be modal. But does causation count as modal or non-modal? Are propositions? (Many take them to be sets of possible worlds, which would seem to make them modal). What about resemblance? It appears non-modal, but on Lewis’s account, for example, resemblance between two things is a matter of both being members of a class of objects which includes mere possibilia among its members. Is colour a modal notion? After all, dispositions are often thought to play a role in its explication. Is mathematics modal? Whether a given domain, or set of predicates, or set of concepts, count as modal or not may often be a difficult question. Comparison between different accounts of modality will be most straightforward when what is non-modal is taken to be common ground between them. However, given the pervasiveness of modal notions and the amount of work they can potentially be made to do, comparing two accounts of modality which are not very similar with respect to what they take the non-modal to be can be a difficult business. This might not be terribly surprising, of course: Duhem and Quine have taught us that theories must often be evaluated as complex wholes, rather than discrete chunk by discrete chunk.25

Hopefully more could be said about the distinction between the modal and the non-modal than the pessimistic conclusion that this distinction only makes sense relative

25 Duhem and Quine might wish to teach us even more holistic lessons than this, I suspect — but I myself resist Quine’s rather extreme holism about confirmation, let alone about meaning.
to a substantial agreement about what is non-modal. Unfortunately, I have as yet been unable to provide a precise and compelling account of this distinction. There is a style of account which is developed on intuitive lines and has some appeal, but which on closer inspection suffers a collapse into the unsatisfactory account of modal primitives which was distinguished above at p 45: that the only systems with modal primitives are those which take the paradigmatically modal operators as primitive (or one of them as primitive — there is no need to treat them both as primitive). I cannot but feel that something like the style of account which can be developed must be correct — so hopefully modification of this approach will yield the desired solution. Let me then present the approach, complete with its refutation. Hopefully others (or a later time-slice of myself) will be able to see a way to improve the account: and at the very least, presentation of this account should make a little clearer what the distinction between modal primitives and non-modal primitives (and, indeed, the distinction between the modal and the non-modal) amounts to, or at least what it amounts to when I employ that distinction.

The account of how to tell what a theory treats as modal and what a theory treats as non-modal is as follows: a given theory (leaving aside eliminativism about modality, which we will concede is not committed to any modal primitives if it is consistent) will employ some paradigmatically modal notions or devices. The best examples of paradigmatically modal notions are those expressed employing the necessity and possibility operators, but the exact boundaries of the paradigmatically modal might be vague. For purposes of this account, let us make the simplifying assumption that only employment of the necessity and possibility operators are to count as paradigmatically modal. There are also presumably going to be other phenomena (please understand "phenomena" in as neutral a way as possible) which can26 be analysed, or reductively explained, or reduced, in terms of the modal operators, either in whole or in part. (Candidates for this treatment might include dispositions, counterfactuals, hyperintensional states such as propositional attitudes, nomic laws, and so on). Let us call these phenomena the modal phenomena (and include the phenomena represented with the modal operators among the modal phenomena as well). Call all of the other phenomena the non-modal phenomena. With this distinction between the modal and the non-modal phenomena, we can provide an account of modal primitives which accords fairly well with the account outlined above, where modal phenomena are identified in terms of those primitives postulated primarily to provide a basis for modality, or those primitives that would not be postulated except to account for the modal. A primitive can be said to be a modal primitive if i) it is a primitive postulated as part of the explanation of the employment of the paradigmatically modal devices (the obvious necessary condition) and ii) it is not employed in the explanation of any non-modal phenomena. Through

---

26 Can by the lights of the theory in question, that is.
meeting the second condition, the primitive is ensured to play a distinctive role in the explanation of some modal phenomena (providing the primitive is not explanatorily useless, and with the caveat that it is not required to play a role in the explanation of itself unless explanation is taken to be reflexive).

The requirement that the phenomena in question can be analysed modally, rather than the requirement that they are or must be so analysed given the theory may not seem obvious. However, if it is required that the phenomena are so analysed (or, a fortiori, that they must be so analysed), the account given faces collapse in any case where the paradigmatically modal phenomena are further analysed. Allow me to digress to explain why.

If modality can be further analysed (say in terms of phenomena which together we can call \( \varphi \)), then any putatively modal phenomena may well be able to be analysed or explained directly in terms of \( \varphi \) rather than employing the paradigmatically modal phenomena in the analysis or explanation. The putatively modal phenomena, if analysed in terms of \( \varphi \) rather than the analysis in terms of paradigmatic modal phenomena, would then count as non-modal, since the paradigmatically modal phenomena would not be invoked in the analysis.

For this collapse to go through, analysis must meet a certain condition (let me dub it the total analysis transitivity condition, or TAT):

**TAT:** When a phenomena \( A \) is able to be analysed entirely in terms of phenomenon \( B \), and \( B \) is able to be analysed entirely in terms of phenomenon \( C \), then phenomenon \( A \) is able to be analysed entirely in terms of phenomenon \( C \).

I will not argue for this principle here, since to do so would require a justification of a general picture of analysis of the sort with which I am concerned. However, I believe this principle is a feature of many of the plausible accounts which could be given of analysis, and an account which collapses when added to this principle is very likely no good. To show that this condition is enough for the collapse I described above, let us take a given putatively modal phenomenon \( M \), which can be totally analysed in terms of a paradigmatically modal phenomenon (\( P \)) plus some (possibly empty) set of distinct explanatory resources: let us call these distinct explanatory resources together \( E \). \( M \), by stipulation, can be entirely analysed in terms of \( P+E \). \( P \) is able to be analysed entirely by \( \varphi \), as stated above, so \( \varphi+E \) should be able to explain \( M \). (This does not follow immediately, but does given any of a variety of plausible additional conditions: such as if there is something distinctive which counts as what in \( M \) is not explained by \( E \), or if a limit case of complete explanation is something's being explained by itself (in either case
$P+E$ is presumably able to be explained by $\varphi+E$, or other constraints which serve more directly to capture the idea that the contribution to an explanation which a phenomenon provides can be provided by the phenomena which fully explain that phenomenon.

Since $M$ can be analysed entirely by $\varphi+E$, it is able to be analysed in terms other than modal ones, so if the requirement for $M$ to be modal was that its analysis was, or indeed must be, in terms of $P$ then $M$ would not count as a modal phenomenon after all. This would go through for all putatively modal phenomena other than the paradigmatically modal ones, yielding the result that when the paradigmatically modal phenomena have a further explanation, there are no modal phenomena other than the paradigmatically modal phenomena—a counter-intuitive result which does not yield a very useful characterisation of the distinction between the modal and the non-modal. This proof, which despite its employment of abbreviative symbols is not meant to be rigorous, is nevertheless enough of a reason to suppose that the analysis should not be in terms of what is or must be the analysis of a given phenomena, but only in terms of whether it can be analysed modally. After all, the result that there are no modal phenomena besides the paradigmatically modal when the paradigmatically modal phenomena are not themselves primitive is counter-intuitive and robs the account of much of its usefulness. So let us take it that the possibility of analysis at least partly in terms of the paradigmatically modal is the mark of a phenomenon’s being modal, and return from this digression.

An account of modal primitives given in this form allows that a non-eliminativist about modality might account for the modal in terms of the non-modal: for if the materials which the realist about modality employs to explain the modal are materials that s/he was required to employ in any case to account for non-modal phenomena, then the definition in question gives the result that the theory is not committed to modal primitives, which is as it should be. An account of this sort also allows that what phenomena counts as modal might vary from theory to theory, which is also plausible. (It does not necessarily allow that what really are the modal phenomena might be theory-relative—for the fact that a given theory disagrees with the (a?) correct theory of the world might just show that the disagreeing theory is false insofar as what is modal by its lights differs from what is modal by the correct theory’s lights.) This account seems to have the features we would want of an account of how to tell what the modal primitives of a given theory are. As I mentioned above, however, this account suffers from a defect that renders it almost worthless as it currently stands.

When the necessity operator or possibility operator (or both) are taken to be primitive, this account delivers the results that one would expect, and which are intuitively correct. Even when the account is formulated so that the modal is defined in terms of what can be analysed in terms of modality, however, the case where the paradigmatically modal phenomena are not taken to be primitive still yields unacceptable
results — results so unacceptable, in my judgement, that the account cannot be satisfactory. Consider again the case where the paradigmatically modal phenomena are analysed in terms of something (or things) further, φ. Let us consider φ itself. Being more fundamental than the paradigmatically modal phenomena, it will not be analysed in terms of them. Therefore, according to this account, φ is non-modal. Now, φ will either consist of one primitive, or several primitives, or be analysable further until it is analysed in terms of primitives alone. If it can be analysed further, then the primitives into which it can be analysed play a role in the explanation of a non-modal phenomenon (φ). So those primitives are not modal primitives by the light of the account given, and since the only primitives employed in the analysis or explanation of the paradigmatically modal are those which explain φ, there will be no modal primitives in the system.

If, on the other hand, φ consists only of primitives, the case needs to be stated a little differently. If φ consists of more than one primitive phenomena, then those primitive phenomena each play a role in the explanation of something non-modal — namely, the phenomenon of those phenomena holding, or occurring, or existing (or whatever the suitable verb is) jointly (i.e. φ). However, if φ is a single primitive, explanation or analysis needs to be reflexive for it to follow directly that φ forms part of the explanation of a non-modal phenomenon (otherwise we merely get the unusual conclusion that φ, a non-modal phenomenon, is also a modal primitive). However, if there is even a single primitive in the system other than the one which constitutes φ, an argument can still be developed to show that φ is a non-modal phenomenon. For consider the phenomenon which is the joint holding (occurring, existing) of φ and another primitive. It cannot be modal, since it is explained in terms only of two primitives, neither of which is paradigmatically modal (since we are supposing that the paradigmatically modal phenomena are not primitive). And φ is part of the explanation of that phenomenon. So φ is part of the explanation of a non-modal phenomenon, so is not a modal primitive.

It is unacceptable that this account delivers us the conclusion that there are no modal primitives unless the paradigmatically modal phenomena are taken to be primitive. At least, it is unacceptable to me: perhaps it can be shown that this is the conclusion that is reached from any reasonable account of which primitives are modal primitives, but I very much doubt it. What needs to be added is a further account of the difference between modal and non-modal phenomena — perhaps the only phenomena which we
should classify for purposes of deciding which primitives are employed only to explain modal phenomena are those which we are faced with at the start of our explanatory project, so that phenomena postulated in order to best explain our starting point are only assigned their classification derivatively (e.g. they are modal if they are explained in terms of a primitive which is not employed to explain any of the starting phenomena which cannot be explained in modal terms). This itself might need to be refined — it is rather vague which phenomena we are confronted with pre-theoretically and which are only theoretical postulates, and the contingent vagaries of the theory-construction process might not be the place to look for a principled metaphysical distinction like that between the modal and the non-modal. Hopefully, however, some refinement along these lines can be made to work, or perhaps a variation at some other point in the account. Another point at which the account might be faulted, for example, is that I have been rather liberal with my conception of phenomena so that, e.g. the holding jointly of phenomena (or composition of several phenomena, or conjunction of phenomena, or however it is to be put) is itself a phenomenon. Perhaps a sparser conception of phenomena might be employed so that it is harder to construct phenomena explained by primitives which explain paradigmatically modal phenomena. And of course there are undoubtedly other routes which might be explored. I will not attempt such explorations here, and will make no further attempt in this work to develop a general account of the distinction between the modal and the non-modal, or of the distinction between modal primitives and others. I shall instead assume both of these distinctions are coherent and useful, and employ them on the basis of optimism that they can be properly accounted for eventually.

Before I leave this topic, there is one other approach that I should briefly mention, if only to show that it does not make too much difference in this context. One might wish to define the modal in terms of whether or not a phenomenon is analysed in terms of possible worlds rather than whether it is analysed in terms of facts expressed employing modal operators. In terms of the account given, this could be captured by taking the "paradigmatically modal" phenomena to include, or maybe even entirely consist in, possible worlds. The same problems will arise in dividing the modal and the non-modal, and in particular the division of primitives into the modal primitives and the others will still cause problems when the possible worlds are not themselves taken to be primitive. This is part of a more general point — additions or variations to the class of what is taken to be paradigmatically modal will not rescue this analysis in general, nor will they be inconsistent with it: it will just be a matter of "plugging in" different phenomena as the paradigmatically modal phenomena in the appropriate place in the theory. I will often take possible worlds to be modal phenomena in this thesis: though whether they are so because they are paradigmatically modal, or because they are analysable (wholly or partially) in terms of the phenomena captured by the use of modal operators, or for some other reason will often not be relevant.
This chapter by no means provides a complete account of the connections between the modal and the non-modal or between modal operators and possible worlds, let alone anything like a comprehensive account of how one should go about connecting theories to each other or evaluate theories for theoretical commitments. It does, however, at least serve to make a start by mentioning the salient options for explanation in such metaphysical cases and clarifying some of the features of the concept of theoretical primitiveness — a concept which is often employed, but very rarely explained. This concludes the second of the two “conceptual exploration” chapters of this thesis: the next chapter continues with a theme related to the question of the primacy of modality vs possible worlds, by examining the so called “Hazen cases”: a group of modal expressions which supposedly indicate the naturalness and attractiveness of taking possible worlds to be primary.
Chapter Three — Expressive Completeness Without Possible Worlds: The Hazen Cases

Introduction

Much has been made of six examples produced by Allen Hazen (Hazen 1976) of statements which are modal in character but which, according to Hazen, defy translation into the modal language of box, diamond and arrow. However, according to Hazen, they are susceptible to fairly obvious translations into the modal logic of quantification over possible worlds. This, claims Hazen, is an argument for analysing modal talk in terms of possible worlds, and indeed for taking our ordinary modal discourse as being disguised discussion of possible worlds and their relations.27

The obvious response to this sort of point is to say that while quantified modal logic (hereafter QML) might lack the resources to deal with some special cases that can be handled by quantification over possible worlds (hereafter PWL, for “possible worlds logic”), this is only a problem with QML as it presently stands, and it does not mean that there is any problem in principle with translating our modal talk in terms of a handful of modal operators which might not be themselves further analysed into more conceptually basic operators. The immediate response to some cases raised by Hazen and Lewis might be to seek to augment QML with one or two other modal operators — most famously, for example, an “actuality operator”, or some such.

Unfortunately, small alterations like this have not been good enough to deal with the full range of cases that Hazen offers. Attempts to deal with the most puzzling of such cases by adding more modal operators lead to unpleasant results like the need to postulate an infinite number of modal operators, as for example Graeme Forbes does (Forbes, 1989, pp 86-88).28 Quite apart from the suspicion that introducing an infinite number of operators in such a way as to make them apparently mimic existential and universal quantification is just a matter of dishonestly smuggling quantification over worlds back into modal logic (a charge which Forbes attempts to rebut), it would be nice if we could rely on a few comparatively simple modal operators to give us our QML. For this reason, I will seek to provide in this chapter translations of the natural-language modal claims offered by Hazen into a QML which needs no modal operators other than the operators of necessity and possibility. As well as the question of whether or not we are able to represent these natural language claims without quantifying over possible worlds,

there is also the question of whether it might be the case that capturing the content of the claims employing modal operators is more perspicuous than employing quantification over worlds: this speaks to the issue of whether or not ordinary modal language is disguised quantification over possible worlds or not. This issue is also important to the dispute between "possible worlds first" approaches and "modal operator first" approaches, since if it turns out that our use of modal operators has been implicitly quantifying over possible worlds all along this gives us *prima facie* reason for supposing that our regimented theories should also treat the idiom of possible worlds as the basic one. The final issue with which I will deal is the acceptability of the non-modal resources I am employing besides the ones which standard QML and PWL have in common: it would be better if the non-modal resources employed were independently attractive, so that they did not count as a cost for the QML option as against its PWL rival.

1. Neutral Quantification

One of the resources which I will employ to provide the promised QML translations will be a resource which may be disliked by some defenders of possible worlds. Nevertheless, this does not make the analyses proposed worthless: at worst it only makes them controversial. The point is to construct a quantified modal logic alternative to the possible-worlds translations of some portions of natural language: it is a separate issue whether this alternative is a metaphysically preferable one to the possible-worlds analysis. In addition, after I have shown that the resource provides the means to translate the first three Hazen cases I will argue that there are independent reasons for employing this additional resource: reasons that hold even for those who already invoke quantification over possible worlds in their translations. Furthermore, I shall demonstrate that some of these independent applications are *non-modal* applications (in at least the intuitive sense of non-modal — we saw in the previous chapter how difficult it is to precisely articulate the division between the modal and the non-modal).

The controversial resource I will employ is the following: I will quantify over things that do not exist. I am not prepared to say that they exist, of course, but for all that, I am prepared to say that some things do not exist. If pressed for an example, Pegasus, or phlogiston, or Zeus or The Absolute are all quite serviceable. Take these examples together, and they should show that at least four things do not exist! Other reasons for thinking quantification over non-existent objects is acceptable (*pace* many free logicians, as well as their classical cousins) are considerations such as that it seems to be

---

28 Though the credit for pioneering the approach defended by Forbes is Peacocke's (see Peacocke 1978).
true that all flying horses do not exist, that most Greek gods were worshipped, that it is possible that there be some purple elephants, even though there are none.

I will allow for objects which do not exist in the logic in a way fairly similar to the standard manner for the so-called "ontologically neutral" logics, such as that of Routley 1980. First, a terminological note: since "\( \forall \)" and "\( \exists \)" are so often taken to be existentially loaded (after all, "\( \exists \)" is even called the "existential" quantifier), I will use different notation to express the quantifier for "all" and the quantifier for "some" in what follows: the equivalent of "\( \exists \)" in the logic that follows will be "\( P \)" in honour of the Routlean particular quantifier (Routley 1980) — although it need not behave quite like a Meinongian quantifier either, as we will see. The corresponding equivalent of "\( \forall \)" will be "\( U \)" (to represent the universal quantifier). Next comes the introduction of an existence predicate: \( E \). (After all, once we are quantifying over things that do not exist as well as things that do, we should have something to mark the distinction). To say that a given thing (for instance \( a \)) exists, then, we would write "\( Ea \)". To say that something exists, we would write \( (Px)(Ex) \). To say that all objects exist we would write \( (Ux)(Ex) \), and so on. If we wished we could add conditions to the effect that an object’s being in possession of a "genuine" property or standing in a "genuine" relation entails that the object exists. Such "genuine" properties and relations could include most ordinary extensional properties and relations. If the most controversial claim of the Meinongians (especially contemporary Meinongians such as Routley, 1980) is that non-existent objects have such genuine properties and relations (for instance, that there is a gold mountain that truly is golden, and that Sherlock Holmes truly did live in London), then our neutral quantification does not imply such Meinongianism. While I can allow for a restriction that non-existent objects do not (actually) have any genuine properties (however exactly "genuineness" is to be fleshed out), I will not require that such a restriction be put in place. Nothing hangs on whether the restriction is put in place or not for the purposes to which neutral quantification will be put in this chapter.

The above strategy might seem to carry with it unacceptable metaphysical baggage: after all, quantification over the non-existent has been thought by many to carry metaphysically unacceptable consequences. It might be thought that this renders using such techniques unavailable to sensible modal logicians, and so neutral quantification cannot form part of a defence of Quantified Modal Logic against the claim that PWL is needed to achieve the same expressive powers as modal talk in a natural language. I

---

29 Many defences of neutral quantification have been launched based on the need for such quantification in intentional contexts. I will not comment on this controversy, however, as there would not be room to do it justice in this chapter. Suffice it to say, if neutral quantification were shown to be indispensable for these sorts of reasons, this would be another argument that it is an acceptable logical resource.

30 See most famously Quine 1953.
reject this claim for several reasons. The first is that this is an issue of logic, not
metaphysics: if the argument advanced against QML is that it cannot translate modal talk
without unwanted metaphysical commitments, this is a separate issue from the issue of
whether it can translate the claims at all. Secondly, it is an option Hazen himself seems to
countenance by his discussion of "outer quantification", a style of neutral quantification
which differs technically from the one I have presented and differs in that the "outer
domain" of quantification is limited to the sphere of the possible, whereas the neutral
logic I have proposed quantifies over everything of which it is true that it does not exist.
Thirdly, an *ad hominem* is appropriate here: many people who wish to use PWL do not
in fact want to claim that worlds, or at least the supposed contents of worlds, such as the
object which is a talking donkey, really exist. Rather, they wish to claim that the truth or
acceptability of quantification over such things in PWL is to be explained in some other
manner — it reduces to claims about strings of sentences, or members of the set-theoretic
universe, or claims about universals and their connections, or about abstracta with *sui
generis* modal properties, or whatever. These people include some ersatzers, all
fictionalists, some structural-universals proponents and many other less common views.
In fact, very many PWL proponents (though not the Lewisians, of course) by their own
lights do not oppose quantifying over things which do not exist when it comes to doing
modal logic. Using an ontologically neutral logic to deal with modal cases, then, seems
no more objectionable than the theories of most of the prominent mainstream proponents
of possible worlds and the accompanying PWL.

Finally, it is my belief that the logical and metaphysical issues surrounding
ontological quantification are separable: I believe one can give an account of non-
existentially committing talk as being reducible to existentially committed talk:
metaphysically reducible even if not semantically or logically reducible (in other words,
one can argue that the truthmakers for claims involving non-referring constants or
quantification over non-existent objects are all actual and existent, even if it is not the case
that existentially uncommitted claims in general mean the same as some set of existentially
committed claims). Of course, if there are semantic or logical reductions available as
well, all the better! One might even take ordinary language modal claims which require
quantification over non-existent objects to mostly be false in virtue of the fact that there
are no non-existing objects (in the broader sense where that quantification is not restricted
by meaning alone to the existing): that is, one could have an "error theory" about non-
existent objects and take this to have fairly radical modal implications, without wishing to
forbid *claims* about non-existent objects to be made: after all, the fact that one's
opponents (or folk opinion) hold a false metaphysical doctrine is no reason in general to
judge that one's opponents (or the folk) have fallen into meaninglessness. Since this is a
chapter focussing on the logical issue of expressibility rather than the metaphysical
disputes surrounding the resources required to achieve such expressibility, metaphysical
worries people have about certain accounts of the basis of neutral quantification are somewhat beside the point (though admittedly very interesting in their own right). I will therefore put such worries aside.

The formal semantics for this neutral quantified modal logic is quite straightforward, provided one does not put any special conditions on the existence predicate. The semantics will be quite standard constant domain semantics — that is, the neutral quantifiers will range over the same domain of "objects" regardless of world, and the addition of an existence predicate, that will function like any other predicate. It will thus most likely vary its extension from some worlds to others. The existentially-loaded quantifiers can then be defined in terms of the neutral quantifiers and the existence predicate in the obvious way. Note that such a semantics validates the neutrally quantified versions of the Barcan formulae, but does not produce as theorems the versions of the Barcan formulae which have existentially-loaded quantifiers.

2. The First Three Hazen Cases: Problems of Existence

With the necessary resources in place, let us get down to the first few Hazen cases. These cases can be dealt with easily, as they take advantage of the fact that a logic which engages in quantification over possible worlds (hereafter "possible world logic" or PWL) is better able to deal with the predicate ".. exists" than classical predicate calculus, at least on the surface. However, the neutral logic I have set out above deals with such cases easily. The formal translations go as follows:

case 1: "Something exists necessarily" (Hazen 1976, p 29)
translation: (Px)L(Ex)

case 2: "There are unactualised possibilia" (p 31)

There are several translations of this available — many of these are obscured by the link between possibly existing and existing which Lewis and Hazen make (which is the assertion that, strictly speaking, anything that possibly exists does exist simpliciter). Some are controversial, and depend in part on claims about whether or not the phrase "there are" and its equivalents in natural languages carries existential commitment (as for example Lewis 1990 and Quine 1953 claim it does, and others, e.g. Routley 1980, claim it does not). I take it to be an advantage of the modal logic employed here that it does not

31 In the case of Hazen it may not exactly be that he thinks that every possible object exists: it is just that he employs existentially loaded quantifiers to talk about them in his 1976 paper. He at least talks as if he thinks that every possible object exists: and it is this way of talking which tends to obscure the distinctions my four translations capture.
force one reading of this phrase, given that it is a matter of controversy what its interpretation might be. There is also a genuine dispute about whether “unactualised” means something other than “does not exist” (e.g. “does not exist in the space-time connected to here and now”, as Lewis suggests).

1st reading (which does not assume that “there is” has existential import, and that unactualised means “does not exist”): \((P(x)(¬E(x) & M(E(x))))\)

2nd reading (which assumes that “there is” does have existential import, and that unactualised means “does not exist”): \((P(x)(E(x) & ¬E(x) & M(E(x))))\) — something obviously false, but a reader who takes the assertion this way does take the assertion to say something false.

3rd reading (which does not assume that “there is” has existential import, and that “.. is actual” and related predicates mean something different from the various predicates which are used to make claims about existence): \((P(x)(¬A(x) & M(E(x))))\)

4th reading (which assumes that “there is” has existential import, and that “.. is actual” and related predicates mean something different from the various predicates which are used to make claims about existence): \((P(x)(E(x) & ¬A(x) & M(E(x))))\)

case 3: “One object is ontologically prior to a second just in case it is not possible for the second to exist without the first one’s existing as well” (p 32)

\(x \text{OP} y = \text{def} L(¬E(x) ⊃ ¬E(y))\)

All three of these cases seem to rely for their bite on the rather unfortunate feature of classical logic which is its inability to deal with the predicate “.. exists” properly. This is a serious flaw, indeed, but it is a flaw that is not directly tied up with modal issues — no matter what one’s position on modality, one still has the problem of saying things of the form “a does not exist”, where a is a place-holder into which some name or other can be substituted. Lewis’s modal realism disguises this to some extent, because of the staggering number and variety of things he takes to exist, for instance, the older sister I never had exists (strictly speaking, I have a counterpart which has an older sister, and that woman exists — but when we are speaking with the vulgar, this means that she is an older sister I could have had, but did not). In a pluriverse containing golden mountains, flying horses named Pegasus, blue swans, and much more besides, we run into the problem of having to talk of “what is not” far less often.
There are no obvious equivalents to such claims in the language of a first-order PWL, as these are the sorts of claims that we would only expect to be able to translate in a second order fashion. If PWL is to be preferred to first-order QML on the grounds of expressive power, then it seems that second-order QML must be preferred to PWL on those same grounds. There are some rescue manoeuvres that the defender of PWL may wish to attempt here, but I will argue that they are unsuccessful, using (1) (and (1')) as my example. The first is an attempt get around the problem by insisting on the second-order quantifiers be restricted in a certain way so as to make them more analogous to quantification over possible worlds, and the second is an appeal to mereological devices.

The first rescue attempt is something like this: analogously to the neutral logicians who wished to quantify over only possible objects, there will be neutral logicians who will want to quantify only over possible collections and possible pluralities. If, furthermore, it is believed that a collection or plurality exists only if all of its members exists, then such a logician will interpret the two examples I offered in a fairly restricted manner. Without this assumption, there is no need to suppose that all the members of the relevant collections or pluralities are co-possible: whereas once this assumption is made, the claims become equivalent to the following:

(1*) There are two collections of co-possible objects (or: there are some co-possible objects and some other co-possible objects) such that necessarily if all of the members of the first (or: all of the first) exist, none of the members of the second do (or: none of the second do), and vice versa.

and

(2*) There is a collection of co-possible objects (or: there are some co-possible objects), each of which is co-possible with one (indeed, all) of the others, and necessarily one of them exists.

Note that the second phrase of the above is pretty much redundant — I put it in only to preserve the similarity to (2).

The restriction of collections and pluralities to co-possible ones is not one I endorse. For one thing, it looks like we just do talk about all the possibilia plurally, and do talk about classes of possibilia that could not co-exist. However, if it is granted for the sake of the argument, it might seem that a defender of PWL has a better chance to translate such claims. For a start, the second claim will have as a PWL paraphrase the

---

42 Many defenders of possible worlds would not be happy with this line of response, however: for
second of Hazen’s challenging cases — as if that PWL claim is true, there will be a collection satisfying (2) (namely, the collection of all the objects in the world in question), and if there is a collection of co-possible objects that are such that necessarily one of them exists, then those objects must all exist in a particular world (otherwise they would not all be co-possible), and that world would have to satisfy Hazen’s formula. Note that without this restriction to co-possible collections, however, (2) could be satisfied while Hazen’s second challenging formula was not: just consider a 4 object, 3 world model where A and B exist in the first world, B and C in the second, and C and D in the third. (2) is satisfied by {A, B, C, D} even though Hazen’s formula is not.

It might be thought that (1*) is susceptible to a PWL paraphrase too, along the following lines (in Hazen’s PWL):

\[(\exists w)(\exists w')(\forall w''') (\forall x)(\forall y)(((Ixw \supset Ixw''')) \supset (Iyw' \supset -Iyw''')) & (\forall z)(\forall z')(((Izw' \supset Izw''') \supset (Iz'w \supset -Iz'w'''))
\]

(There are two worlds such that for any world, if all the objects in the first are in that world then no objects from the second are, and if all the objects from the second are in that world then no objects from the first are.)

However, this is not an adequate paraphrase: for it may be the case that no two worlds satisfy the above condition, and yet there be two collections such that it is not possible that any of one could co-exist with all of the others. For suppose that there was some object that necessarily existed: then no two worlds would satisfy the above PWL formula. There being a necessary existent would not prevent there being collections conforming to the required condition though: for unlike worlds, arbitrary collections of objects are not required to contain all necessary existents. Even if there were no necessary existents, it would still be easy enough to construct models where no worlds satisfied the PWL formula offered as a paraphrase, and yet the claim about there being two collections obeying the QML formula would be true. The problem with PWL, to cast it in terms of quantifying over collections of possibilia, is that in general PWL only has devices to simulate quantification over collections of co-possible objects, and even then only when it is possible that those objects in the collections be all the objects that exist (i.e. where it is the case that those objects are all the objects in a world). So it can be seen that making the assumption that quantification over collections is only over collections of things that are co-possible does not solve the PWL supporter’s difficulty, since the collection of members of a possible world are not just co-possible, but it must also be possible for them to be all the objects.

example, those who follow Lewis’s account of properties and relations need to insist that collections and/or pluralities exist even when all of their members cannot be found at the one world.
Next, it might be thought that (1*) can be translated adequately into a PWL if mereological devices are added to a PWL: for then quantification over parts of worlds can be used to pick co-possible objects without picking out all the co-possible objects in a world, as necessary existents or other such inconveniences could be got around simply by mereological subtraction. But even this will not work. Consider, for example claim number (1). Suppose that this were true in a two-world model because one world contained the Ship of Theseus and the other contained a ship which had say, three quarters of its planks the same as the Ship of Theseus and the remaining quarter were different (call it Ship of Theseus*). Now while it is true that the collections {Ship of Theseus} and {Ship of Theseus*} conform to condition (1), just as obviously the mereological sums, the first of which contains the Ship of Theseus and the other which contains the Ship of Theseus*, do not obey (1#), where (1#) is

(1#) There are two mereological sums of co-possible objects such that necessarily if all of the parts of the first exist, none of the parts of the second do, and vice versa.

This is because, of course, three-quarters of the planks in the sum of the Ship of Theseus will exist whenever the mereological sum of the Ship of Theseus* does. (1#) might not even be satisfied at these two worlds at all, if indeed the worlds contain the same planks, nails etc., but it is just the case that they are arranged differently. So the truth conditions for (1) cannot be captured by first-order PWL, even with the co-existence restriction, and even with the addition of mereological devices.

So my example (1) survives a variety of attempts to render it in first-order PWL, even with the restriction that all the members of a set or plurality of possibilia must possibly co-exist. (1) does, however, rely heavily on neutral quantification — it says that there are two such collections... I have defended the propriety of employing this device already, but since the device is controversial, it is worth noting that it does not need to be relied on as heavily as I have done up to this point. For a statement which amounts to (1) with the restriction that the members of each collection must possibly co-exist can be stated as follows:

as a claim about collections (the “E” is for existentially loaded quantifiers):

(1E) It is possible that there be a collection of co-possible objects, it is possible that there be another collection of co-possible objects, such if all the members of the first set exist then none of the second do, and vice versa.
\[ M(\exists c) M((\forall k)(k \in c \supset (\exists m)(k=m))) \land M((\exists c') M((\forall n)(n \in c' \supset (\exists r)(n=r))) \land L(((\forall s)(s \in c \supset (\exists t)(s=t))) \supset (\forall u)(u \in c' \supset (\forall v)(v \supset (\forall w)(w \in c' \supset (\exists x)(w=x))) \supset (\forall y)(y \in c \supset (\forall z)(z \supset (y=z)))))) \]

and as a claim about pluralities:

It is possible that there be some co-possible things, and it is possible that there be some other co-possible things, such that if all of the members of the first set exist then none of the second do, and vice versa.

\[ M(\exists X)L(\forall h)(MXh \supset LXh) \land M((\forall j)(Xj \supset (\exists k)(j=k))) \land M(\exists Y)L(\forall m)(MXm \supset LXm) \land M((\forall n)(Yn \supset (\exists r)(n=r))) \land L(((\forall s)(Xs \supset (\exists t)(s=t))) \supset (\forall u)(Yu \supset (\forall v)(v \supset (\forall w)(w \supset (\exists x)(w=x))) \supset (\forall y)(Xy \supset (\forall z)(z \supset (y=z)))))) \]

Of course, even employing existentially loaded quantifiers like the translations above will not settle all the qualms of some who insist on so called “actualist” quantifiers (in the terminology of Fine in Prior and Fine 1977, among others). For the above translations to capture the intended meaning properly, objects have to be able to be members of collections or belong to pluralities even when not all the members of the collection or plurality exist. Since I am also assuming rigidity of collection membership and plurality membership, this means that a collection or plurality could have an object as a member even when that object did not exist. Membership, then, and the relationship of belonging to a plurality, will still not be “genuine” relations, since they can hold between existents (collections or pluralities) and objects which do not exist. Alternatively, if it is thought that a collection does not exist unless all of its members do, the above translations need to be understood so that an object can be a member of a set even when both do not exist (which would also make the relation a non-genuine one). More could be done perhaps to accommodate the worries of the quantifier actualist: perhaps collection membership or plurality membership could be made non-rigid, but in such a way that there was a maximum membership of a given collection (or plurality), but the collection (or plurality) could have as members some but not all of its maximum membership at worlds where only some of the objects in its maximum membership existed (say, the extension of a collection or plurality at worlds which have fewer members existing have just the existing objects which can also be found in its maximum membership as its members). In a sense though whether these qualms can be accommodated is a side issue to the point of this chapter, since a defender of QML need not feel these qualms when presenting translations of the modal claims at issue. And the defender of PWL will typically have even less excuse to invoke such qualms against the defender of QML, since
PWL almost invariably has quantifiers which are not bound to the domain of a single world. Even without the full resources of neutral quantification, (1) shows that second-order QML can translate ordinary(-ish) language modal claims which a standard PWL such as Hazen's cannot.

There will be cases of apparently second-order modal claims that can be translated by PWL other than claims which can be interpreted as being about collections which have co-possible members which are such that they be all the existing objects. The most significant of such cases will be those that are entailed by claims which quantify over only single-membered collections, or 2-membered collections, or in general, claims about n-membered collections, for some specific n. For instance, the claim that "there are two collections, neither of which has any members which are copossible with any of the members of the other" will be true iff there are two objects which are not co-possible. For if there are no non-co-possible objects, then obviously the claim is false, and if there are just two non-co-possible objects, then the unit collections of each will satisfy the claim. And if there are collections of greater than one member which satisfy the claim, then there will have to be unit-sub-collections of each that do. Since PWL can, at least most of the time, handle claims about the modal properties of individual objects as well as QML, PWL can formalise a claim equivalent to a claim about the members of unit sub-collections just by talking about the individuals which are the unit-subclasses' members. So for example, the PWL formalisation of the claim equivalent to "there are two collections, neither of which has any members which are copossible with any of the members of the other" is simply

$$(\exists x)(\exists y)(\forall w)((Ixw \rightarrow \neg Iyw) \land (Iyw \rightarrow \neg Ixw))$$

(There are two objects such that for all worlds, if one is in the world the other is not.)

PWL translations can also be given in the obvious ways for claims about the modal properties of members of collections when those claims imply and are implied by claims about two membered collections, or three membered collections, or again n-membered collections for any specific n. This is what is to be expected, however: the extent to which PWL can model statements about collections in the special case when statements about such collections implies the truth of claims about some specific number of objects is just the extent to which more generally claims about collections of objects can be paraphrased with first-order formal renderings without having to invoke the machinery of second-order logic. After all, it is well known that when a claim about a collection of objects implies and is implied by some 1st order claim about some specific number of objects, first order logic can translate such claims using the device of non-identity. This type of case of PWL being able to adequately translate claims about modal
properties of collections is thus neither here nor there, especially as 1st order QML will share this degree of expressive power about collections.

What is becoming obvious is that 1st order PWL can only handle a subset of those natural language modal claims formalisable in 2nd order QML. 1st order PWL can only handle the ones which either have a first order QML translation, or are such that require only first order QML plus the ability to quantify over the sets or pluralities which correspond to collections (or pluralities) containing (or being) objects such that they are all co-possible and such that it is possible that they be all the objects. In fact, were we to construct a logic which was like first order QML with the addition that such collections fell under the range of its quantifiers, and the addition of the set-membership relation, then we would obtain a logic with, in effect, the same expressive power as PWL when it came to translating ordinary language modal claims about collections. From this broader perspective, the following reasons can be seen as to why the move to second order translations is not an *ad hoc* response to the Hazen cases, or a simple admission of the superiority of PWL in the field of expressibility: second order QML translations provide an increase in expressibility of natural language modal claims, and they have some grounds in the conceptual analysis of the English version of the modal claims made by the two difficult Hazen cases (as such English claims have natural second-order interpretations), and its invocation provides the means for a general understanding of why PWL has greater expressibility than first order QML in the type of case which Hazen offers.

That 2nd order QML has an expressive advantage in regard to translating natural language modal claims has already been shown. Furthermore, we can see that it achieves its superiority in expressive power without introducing a special realm of objects or predicates (worlds or the "is a world" predicate) or new and unfamiliar relations (such as the relation of "being in a given world" and, depending on which version of PWL is adopted, "is a counterpart of"). True, new resources are introduced (either the relation of set/class/collection membership or the device of plural quantification), but these are obviously resources which are required outside modal contexts for other purposes. The resources it draws upon are only the ones that are naturally suggested by the English renderings of the challenging Hazen cases (as the renderings suggest talk of collections or pluralities), in contrast to PWL which employs introduced specially introduced predicates, so 2nd order QML has a claim to be considered a more natural rendering.

Finally, to show that bringing 2nd order QML into the picture helps to explain first-order PWL's advantage over 1st order QML, we need only consider that the expressive power of PWL in modal cases is equivalent to 1st order QML augmented by devices for quantifying over a special sort of collection (or a special sort of plurality of things) — the collections (or pluralities) which are made up of co-possible objects capable of being all the existing objects (i.e. all the possible collections of world-constituents). It
is not an innate superiority of quantification over worlds, but rather a predictable consequence of the introduction of PWL's novel 1-place predicate and 2-place predicate that PWL should handle the challenging cases but that 1st order QML would not.

"What about second order PWL?", it might be asked. And it is true enough, as far as I can tell, that second order PWL has at least the same expressive power as second-order QML. However, I suspect that second-order PWL has no greater expressive power when it comes to ordinary language modal claims of the sort considered by Hazen and this chapter. Furthermore, second-order PWL, while it may dispense with modal operators, will need to invoke new objects/predicates and relations which QML, even second-order QML, need not concern itself with. While it is true that expressive completeness (relative to natural language modal claims) can be achieved by PWL (providing it is second-order), nevertheless, for all that Hazen's examples show, expressive completeness can still be achieved by (neutral, second-order) QML. Hazen's argument that quantification over possible worlds is needed to formalise the full range of ordinary language modal claims clearly fails, and his suggestion that quantification over possible worlds is to be preferred in general to QML is not adequately supported by the examples he provides. Thus Hazen's cases do not obviously show that there is reason to analyse ordinary modal discourse as disguised quantification over possible worlds. If PWL is supposed to have any advantages over old-fashioned operator-modal logic which are meant to motivate the analysis of modal discourse as possible-worlds discourse then these advantages must be sought elsewhere.43

---

43 An earlier version of this chapter was presented as a paper. Thanks to Greg Restall, Peter Menzies and Timothy Williamson for helpful comments and advice in preparing this chapter.
Chapter 4 — Problems For ‘Strong’ Modal Fictionalism

Introduction

In this chapter and the next, I will shift my attention to issues concerning specific accounts of possible worlds: in this chapter I outline some problems which face modal fictionalism, a variety of theory of possible worlds which is gaining in popularity (or at least notoriety), and in the next chapter I deal with a problem for modal fictionalism and an abstractionist account of possible worlds known as linguistic ersatzism (after Lewis’s dubbing of it in Lewis 1986c).

There are a cluster of problems for modal fictionalism which centre around whether or not it is self-defeating, or nearly so. The ‘Brock/Rosen’ objection (Brock 1993, Rosen 1993) is the best known of these objections, and another which has recently been proposed is an objection of Bob Hale’s (Hale 1995). I will not deal with this sort of problem in this chapter (or in this thesis), since I wish to concentrate on another set of problems. These problems arise specifically for a particular variety of modal fictionalism, and I will argue that each of these problems constitute a reason for abandoning at least this version of modal fictionalism as unworkable45. The exploration of these problems, and what they show about modal fictionalism, is the topic of this chapter.

Not all versions of modal fictionalism: for I will be distinguishing three main varieties of modal fictionalism, and I will only be arguing that the four problems I raise are problems for one of the types of theory which employ the notion of fiction to explain some of our talk about modality, and particularly our talk about possible worlds. I will argue that the other positions, on the other hand, are much better placed to deal with the problems I shall discuss. However, even though I am only attacking only one specific variety of modal fictionalism, I do not take myself to be attacking a straw dummy either: while I will explain that there are genuine modal fictionalist alternatives to the type of theory I am attacking, I take it that the two most prominent modal fictionalisms that have appeared in the literature are versions of the type of modal fictionalism I will be attacking. These two theories are the modal fictionalism presented in Gideon Rosen’s paper named, not surprisingly, “Modal Fictionalism” (Rosen 1990), and the modal fictionalist

---

44 An earlier version of this chapter was presented as a paper at the Australasian Association of Philosophy conference in 1994, and I thank everyone who commented. Thanks also to Cathy Legg, Graham Oppy and Philip Pettit for discussion and comments on various drafts of this paper, and especial thanks to Peter Menzies and David Armstrong for comments and discussion.

45 Arguably the Brock/Rosen objection only has bite against this version of modal fictionalism also: see Nolan & O’Leary-Hawthorne 1996.
approach which was taken in D.M. Armstrong’s *A Combinatorial Theory of Possibility* (Armstrong 1989). It is not clear to me that either of these authors still holds the sort of modal fictionalist position they once argued for (or that Rosen ever held a modal fictionalist position, as opposed to his having merely suggested it). But a theory does not really need current defenders for it to be worth discussing, and I suspect that if modal fictionalism of the type I wish to object to is not dealt with, philosophers supporting it will reappear soon enough. The first thing I will need to do in this chapter though, before I begin discussing the problems, is explain what I am talking about — that is, what modal fictionalism is, the different kinds there are, and more specifically the nature of the version of modal fictionalism that faces the problems mentioned in the title.

The essence of modal fictionalism is the claim that possible worlds do not literally exist (that is, there are no non-actual states of affairs, or no non-actual situations), but that when we engage in talk about possible worlds we are making claims which are literally false (at least when we discuss non-actual worlds) but which are true *according to a story we tell*, and that this story is useful for our treatment of modality. This “truth in a story” is to be understood in the same way as we understand claims like “According to the *Lord of the Rings*, it is true that there are many orcs in Middle Earth”, or “According to the *Monadology* by Leibniz, it is true that there are many monads”. Of course, Leibniz himself did not take the *Monadology* to be merely a story — he took it to be a sober presentation of the truth. Nevertheless, I take it, his views about the existence of Monads were literally false. Analogously, modal fictionalists do not claim that possible world talk is always *intended* to be merely a story by those who use it — that would be absurd. The claim is rather that in fact (nearly all) possible world talk is literally false, but also that some claims about possible worlds are true according the possible-worlds story, and some are not. A sample claim about possible worlds that would count as being part of the possible worlds story, and is true according to the story is:

“There is a possible world where swans are blue”,

and a sample claim about possible worlds that is not true according to the story is:

“There is a possible world where there is a round square cupola on Berkeley college”.

These are both fairly hackneyed examples, of course, but they do serve to illustrate the sorts of claims which are in the story and the sort that are not. Exactly where the boundary is, and why it is there rather than anywhere else, is an interesting question, but it is one that will be answered differently according to different modal fictions.
The general picture offered by modal fictionalism looks quite attractive. If we are not going to be concretists about possible worlds (that is, say that every possible world really exists in just the same way that this one does), and we think that objects like sets of sentences or sets of propositions or bits of our heads or weird abstract selection-functions are just are not the sorts of things that can do the work of possible worlds — in other words, if we also reject abstractionism about possible worlds — then we probably want to say that possible worlds (other than the actual one) do not exist. And given that we are tempted to say that, fictionalism about these possible worlds is the only obvious alternative that does not have a Meinongian tinge. Now none of that was meant to be an argument against concretism, abstractionism or even Meinongianism, but given that as a matter of fact many philosophers find these options unattractive, one can see why they would want to embrace fictionalism about possible worlds.

Suppose then we were to embrace fictionalism about possible worlds. The question would arise as to what implications this would have for our theory of modal claims. After all, if we decide that the claim “There is a possible world where swans are blue” is literally false, what are we to make of the claim “Possibly, swans are blue”? After all, didn’t we take in with our modal mother’s milk the biconditional

“There is a possible world where swans are blue iff Possibly, swans are blue.”?

Does fictionalism about possible worlds mean that we have to deny that there could have been blue swans? It is on the issue of the connection between possible worlds talk and the truth of modal claims that I wish to distinguish my three versions of modal fictionalism.

1. Three Varieties of Modal Fictionalism: Broad, Timid and Strong

The first modal fictionalism I will consider recommends a fictional approach not only to possible worlds, but to modality as well. Since it is fictionalist about both, I will dub this version “broad modal fictionalism”. This sort of position would be very similar to an eliminativist position on modality. For instance, suppose that you thought that there just was nothing in objective reality that corresponded to any necessities or non-actual possibilities. Nevertheless, you might still think that talk about what was possible or necessary was a useful device, or shorthand for dealing with, for example, causation or deductive reasoning. The fiction of possible worlds would then be seen as a handy extension of the basic fiction of modality: not only would one sometimes ascribe the merely fictional properties of “necessity” or “possibility” to statements, or “essentiality” or “accidentality” to properties, one could also use talk of non-existent worlds as a device
to help manipulate these necessities and possibilities. The biconditional could remain perfectly intact as a handy translation procedure from one fiction to another.

It is clear that the above position has a claim to be called "modal fictionalism", for it is the most full-blooded variety of fictionalism about things modal. I do not happen to be in sympathy with it: in fact, I am pretty sceptical about most attempts to be instrumentalist about everyday areas of discourse in this way. However, it is one of the varieties of modal fictionalism that I will not address in this chapter. To rebut this position would be to mount a defence of the basic claim that there is an objective truth which our modal discourse attempts to track — a task which is, while vital, beyond the scope of this thesis.

The other two varieties of modal fictionalism have in common the claim that normal modal statements are objectively true or false, and furthermore that some statements about the possibility of the non-actual or statements about things being necessary are in fact true, and not just true according to a fiction. Where they differ is with respect to the question of priority: are the statements about the facts of modality somehow true in virtue of the truth of the statements about possible worlds, or is it the other way around?

The thesis that the statements about possible worlds (or to be more precise, how possible worlds are described in the fiction) rely on the facts of modality, rather than the other way around, is called by Gideon Rosen a "timid fictionalism" (Rosen 1990, p 354). The main version of fictionalism proposed in his Modal Fictionalism is not this timid fictionalism, but he does not rule out the "timid fictionalist" strategy either. Another expression of this "timid fictionalism" is the modal fictionalism of Hartry Field. In the introduction to Realism, Mathematics and Modality, he explicitly says "in my view, possible worlds are just fictions" (Field 1989, p 41) However, he also takes "logical possibility" as a primitive, and says that he "must accept" that this logical possibility is "not something that must be explained in terms of entities... e.g. ... possible worlds" (Field 1989, p 86). When it comes to logical possibility at least, then, it appears Field is only a timid fictionalist.

I have some sympathy for this so-called timid fictionalism— and furthermore, I believe that a timid fictionalism can be made out which avoids all the problems I will discuss in this chapter, and furthermore can deal with the other problems raised for fictionalism in the literature. The problems I will discuss for modal fictionalism are not meant to be problems for timid modal fictionalism. This leaves us with the other variety of modal fictionalism which takes the truth of modal claims to be derived from claims about possible worlds. This sort of fictionalism, which Rosen calls a stronger reading of modal fictionalism (Rosen 1990, p 354), aspires to be a deflationist theory about modal claims, one which reduces modal claims into non-modal claims (in this case truth in fiction) — or, given that truth according to the fiction might still be some sort of modal
notion, at least reduce many modal notions into one. It aspires to be an "analysis of modal language" (Rosen 1990, p 348) itself, and not a theory merely about our possible-worlds talk.

Rosen employs this stronger understanding for the bulk of his Modal Fictionalism paper (Rosen 1990). Finally, the view expressed in Armstrong’s A Combinatorial Theory of Possibility (Armstrong 1989) was a variety of this strong fictionalism. The strategy was as follows:

> We set up non-existent “merely possible worlds” alongside the actual world, using certain principles. ... “It is possible that p” is then said to be true if and only if a world can be found in which p is true. (Armstrong 1989, p 51)

However, this dependence does not mean that the truth of claims about the merely possible is merely fictional truth: these claims are still literally true on Armstrong’s account (Armstrong 1989, p 50). So this third approach to modal fictionalism, which I have dubbed “strong modal fictionalism” is the approach taken by two of the philosophers most concerned to argue for a fictionalist approach to modality. However, as I said earlier, I think that this “strong” version of modal fictionalism faces some major problems. The first one concerns the artificiality which apparently infects modality on a strong modal fictionalist understanding, while the second two focus on ontological inadequacies which strong modal fictionalism faces, while the final problem concerns its apparent redundancy in an important respect.

2. The First Problem: Artificiality

The “artificiality” worry is that strong modal fictionalism seems to make modality too artificial. Of course, modal fictionalists want to hold that not just any old fiction that talks about possible worlds is the fiction to be used in assessing claims about what is possible. Both Rosen and Armstrong wish to have constrained fictionalisms, and wish to ground these constraints in something actual: Armstrong in basic combinatorial principles plus truths about universals and states of affairs which actually exist, Rosen in basic world-principles which in turn, he suggests (Rosen 1990, p 353), might be justified by appeal to facts about our imaginative practices. In addition, Rosen wishes to include as part of the story of worlds an “encyclopedia” of all truths about the actual world, which is needed, along with the basic principles, to provide a complete fiction of possible worlds (Rosen 1990, p 335).

Let me note in passing that the task of giving a plausible and adequate account of the constraints on selection of the modal fiction is, I suspect, the most difficult task facing a strong modal fictionalist, and a task which it is far from clear that either Armstrong or Rosen have successfully carried out. As an example of the challenge faced, all but the
para-consistent fictionalists will need to postulate non-modal constraints that ensure the selection of a story that only claims consistent worlds exist. But since which non-modal constraints are argued for varies from account to account, this problem will not be further explored: whether a given strong modal fictionalism meets the requirements about constraints will need to be determined by examination of the particular fiction in question. But there is still a sense in which the theory makes modality artificial. As we know, fictions, theories etc. are artificial: we make them up, and they do not exist before we make them up. Leibniz, for example, actually created the Monadology: he did not discover it in the way that, for example, an explorer might be said to discover a new island. Similarly, it is not the case that Conan Doyle found out about Sherlock Holmes by remembering a shadow of the Form of the Detective Story or some such: before Doyle, there were no Sherlock Holmes stories, and after him there were. The Modal Fiction is a fiction like the others — if people had not thought of or considered possible worlds, it would not have existed. So it seems that if people like Leibniz, Saul Kripke and David Lewis had not existed, and nobody else thought up a theory of possible worlds either, then absolutely nothing would have been possible. In such a world, when Aristotle muses about a sea battle, his musings are completely pointless: it would not have been possible either that there be a sea battle tomorrow or that there would fail to be one — and it would not have been necessary that either there was a sea battle or there was not. However important one thinks philosophers might be, one should not think that philosophers are that important. This mind-dependence of possibility should be especially worrying to those who think that causal laws entail counterfactuals, and that counterfactuals involve claims of possibility and necessity — for then causation becomes something that is only in the world because some people told some stories about possible worlds.

There are at least three replies that might be attempted to this artificiality objection, but they seem to me to face troubles greater than the ones they are meant to solve. This first reply consists of admitting that the fiction is artificial but claiming that this does not face the above problems if one adopts a “rigidifying strategy”. The traditional fictionalist translation scheme for modal claims was of the form “Possibly P iff according to the Modal Fiction, P is true at some world”. The fictionalist following the rigidifying strategy will claim that the second half of this biconditional is to be read rigidly, so the biconditional, when spelled out more explicitly, should read “Possibly P iff actually according to the Modal Fiction, P is true at some world”. Since the fiction came to be in the actual world, the truths of modality are actually true. And since the fiction states that there are an infinite number of mind-independent worlds, the fact that in some of these

46 This is in fact controversial: for a discussion of the alternative Platonist theory of fictions, see the last section of this chapter. For now I will assume such Platonism about fictions is false.
worlds life does not arise, or the world ends upon Aristotle’s death, does not imply that in those worlds nothing is possible or necessary — on the contrary, the way fictions like Rosen’s and Armstrong’s actually function, just as much (or almost as much) is possible in those worlds as is at ours. This is because while the fiction does not exist at such worlds, it exists in the actual world, and it is its existence at the actual world from which modal truths are derived. So the counterfactuals like “If people had not come up with this fiction, nothing would have been possible” are not in fact true, and in fact are necessarily false. The fictionalist who responds in this matter may think that the reason why we are worried about artificiality is because the counterfactuals like the ones I have mentioned seem counterintuitive. Once it is realised that these counterfactuals are not true, artificiality should not be something to worry about.

So says the strong fictionalist following the rigidifying strategy. However, this merely makes it a little harder for us to express our worry about the dependence of modality on the fiction. Harder, but not impossible. Davies and Humberstone, in their (Davies & Humberstone 1980), among others, have argued that it makes good non-trivial sense to evaluate conditionals about what follows if the actual world were different. It is still “deeply contingent”, to use Gareth Evans’ phrase (coined in Evans 1979), that anything is actually true according to the Modal Fiction, and we can still object that modality is artificial in a worrying way according to strong modal fictionalism: this is because if there had actually been no Modal Fiction, no statement about contents of worlds would have been true according to it. All the objectionable conditionals can be re-raised as conditionals about what modality would be like if the world had actually been different. This of course relies on having a theory that does not take such “actually-conditionals” to have trivial truth-evaluations, such as a theory employing two-dimensional modal logic for example, and a modal fictionalist might argue that such approaches are fundamentally flawed on other grounds. If a fictionalist is prepared to claim that all counterfactuals with statements about how things actually are to receive trivial truth-evaluations, then this response to the rigidifying strategy will have no bite. Then the worry about the rigidifying response would be the question of how the fictionalist is to justify their claim that their theory is an explanation of modality, since if the fiction were to be different (or actually different) modality would be the same, and in no non-trivial counterfactual sense does modality objectively depend on the fiction. Once we go past the basic material biconditional to attempt an explanation of why the biconditional holds, the rigidifying fictionalist will, I suspect, face great difficulties.

The second response which might be made to the “artificiality” objection is to bite the bullet, and be quite unashamedly anti-realist about modality, and genuinely take the truth of modal claims to be an artifact of our story-telling. While I have little sympathy for this approach to modality, it is one that avoids this objection. However, there does
seem to me to be one obvious problem with this approach — if one were to conclude that all our modal talk was artificial in this sense, then my suspicion is that one should become either a fictionalist of the first sort I described in this chapter: a broad modal fictionalist, or alternatively an out-and-out eliminativist about modality (that is, to say that all modal claims are false), rather than someone who thinks that there are genuine truths of modality, but they reduce to truths about a story we tell.

That may sound question-begging, but to indicate why I think eliminativism is preferable to anti-realist Strong Modal Fictionalism (or that Broad Modal Fictionalism is preferable to anti-realist Strong Modal Fictionalism) let me use an (admittedly controversial) analogy. Suppose that one believed (as many do) that there are no objective, non-artificial positive facts about God. For instance, if one believed that no-one created the universe, that no person sees all and knows all (or even sees-most or knows-most), and so on. In such a case, it seems to me, there are at least two options. One is to become an eliminativist about God (i.e. become an atheist). The other is to say that God exists, all right, but is an artifact of our way of talking, and there are positive truths about God all right, it is just that these are analysed in terms of what is acceptable to say in the God-story. Now, it seems clear to me (and I hope to you as well) that the first way of going is preferable in the God case, and so for the same sorts of reasons I think it is preferable to be an eliminativist, or alternatively a “broad” modal fictionalist, rather than an anti-realist modal fictionalist of this third, “strong” variety.

There is a third response to the “artificiality” objection, which is also a response that accepts that modality is artificial, but is a conceptualism about modality rather than a straight-out anti-realism of the sort just described. Some might not think that this is too different from an anti-realist view, but I think conceptualism and anti-realism can be prised apart. Unfortunately for conceptualism, the only prising apart I can do comes at the cost of making the conceptualist position that is substantially different from the anti-realist position look absurd.

Before I try to untangle anti-realism and conceptualism about modality, I should say a brief word about doxastic and epistemic modalities, and why they are not directly relevant to the conceptualism I am describing. Nearly everyone admits that it is legitimate to speak of such things as epistemic modality (where something is epistemically possible if it is compatible with everything we know), conceptual modality (something is conceptually possible iff it can be conceived of), doxastically possible (where something is doxastically possible iff it can be thought to be true), and so on. Furthermore, it seems clear that we should be somewhat conceptualist about these notions (these modalities essentially rely on facts about our minds and about our conceptual resources). These are not the sorts of modality I am discussing at the moment, however. I am discussing the “objective” modalities, especially logical and metaphysical possibility and necessity, and these are not uncontroversially conceptual.
So, to get back to the discussion, why is a conceptualist strong modal fictionalism not a good response to the artificiality objection? And secondly, what is the practical difference between conceptualism and the anti-realism I discussed earlier? To answer the second question, there are two general varieties of conceptualism which I will outline: “strong mind dependence” and “weak mind dependence”. I will illustrate both approaches in terms of a conceptualism that is more familiar: conceptualism about universals. This conceptualism is one that holds that universals exist, but only in our heads: things fall under the universal “square”, for example, in virtue of our classification system.

The issue which divides strong mind dependence from weak mind dependence is the following issue: are any unseen things square? Or, to put the issue in a more general form, do things which are unconceived of by us have any properties or relations? Strong mind dependent conceptualism will answer in the negative. For a strong-mind-dependence conceptualist things are square, for example, only if we classify them as square. Things which we do not classify (because we have not been exposed to them, for example), must lack properties and relations altogether, because if we do not think of them at all, we certainly do not classify them in one way or another.

This strong mind dependence, it seems to me, is pretty similar to anti-realism of the type I described, and I have no further objection to it to discuss at the moment, except to say that I think someone with this sort of conceptualism should be either an eliminativist or embrace the first sort of modal fictionalism I mentioned, rather than this third variety. However, weak-mind-dependence conceptualism is a substantially different approach, and answers the question about the unperceived and the unconceived a different way. Let us return to universals to examine this new way.

The weak-mind-dependence conceptualist allows that unperceived and unconceived objects have properties and relations. This is because it is not a matter of objects actually being classified that connects them to universals for this sort of conceptualist — it is rather that were an object to be presented to someone, they would classify it thus-and-so. So this sort of conceptualist can allow that some object (e.g. a rock in another galaxy) has the properties of being hard and massive, even though no-one has ever classified it as being hard and massive, because of the fact that were someone to perceive it and think about it, they would (at least ceteris paribus) conclude that it was hard and massive. Weak mind dependence conceptualism can claim that mind-dependence need not lead to the extreme artificiality which strong modal fictionalism has been charged with — there is still room for facts about the world, even granted that the truth in the area of discourse (be it universals or modality) has some essential dependence on our minds and our concepts.

All this might sound attractive (or it might not — I’m no conceptualist myself), but it falls apart when modality is the subject that we are trying to run a weak-mind-
dependence conceptualism about. Suppose that we wanted to say that it was possible that $P$ even though no-one had ever considered $P$, or claimed that it was possible, or claimed that there was a possible world where $P$ held. The weak-mind-dependence conceptualist might say that this is all right, because were we to consider $P$, we would say that $P$ held in some possible world (according to our fiction), or that in worlds where no-one tells the fiction $P$ would still be possible because were someone to tell the modal fiction, $P$ would hold in some world according to the fiction.

However, obviously, the analysis of what made $P$ possible just given explicitly re-introduced modal notions — furthermore, reintroduced modal notions which are just as difficult to deal with as the notion of unthought-of possibilities in the first place. Weak-mind-dependence conceptualism is just not an option when it comes to modality. Here ends my two-part reply to the conceptualist strategy of avoiding the brunt of the charge of artificiality — either the conceptualist reply collapses into something like anti-realism, in which case see my answer to that, or it is plainly incoherent when applied to modal phenomena: depending on which version of conceptualism is being employed. The artificiality objection therefore does not seem to be able to be met satisfactorily (by my lights, in any case) by any of the three responses that I have canvassed.

Before I leave the artificiality objection, let me briefly mention why it afflicts the strong modal fictionalist with more severity than the broad or timid modal fictionalist. Broad modal fictionalists, committed as they are to modality being merely a convenient fiction, seem to avoid the charge altogether: since there really is no modality, the issue of its artificiality does not really arise. And the broad modal fictionalist should be happy to say that the story according to which things are necessary or possible is artificial — that is rather the point of broad modal fictionalism, after all. On the other hand, the timid modal fictionalist is not open to the artificiality charge for almost the opposite reason — the timid modal fictionalist can allow that what the modal fiction says about possible worlds depends on what the modal truth in fact is — and the timid modal fictionalist need have no reason to say that the modal truth itself is at all artificial. The timid modal fictionalist might well agree that the modal fiction is artificial, and would not exist if we had not made it up, and so perhaps even that nothing would have been true according to it if we had not made it up. However, the timid modal fictionalist can just say in those cases that the biconditionals connecting modal truths and what is true according to the fiction do not hold, and this leaves the modality itself untouched.
language modal claims that QML cannot. It might be worth stressing again what is the
issue I am concerned to deal with: I wish to dispute the claim that there is an expressive
superiority when it comes to the translation of modal claims of PWL over QML.

Statements in PWL that are not meant to represent modal claims are neither here nor there.

The above claim has several features which explain why it is not fit for easy
translation into QML. The first is that it talks about “worlds”, and “world”, in the mouth
of a possible-world-logician of the sort that follows the broad outlines of (Lewis 1968),
is meant to be a specific technical term to which nothing corresponds in non-modal
predicate calculus (Wx is a “new... primitive predicate...” according to Lewis 1968). If
we take this at face value, then of course non-modal predicate calculus with the addition
of modal operators will have trouble talking about worlds — but only for the reason that
it has been stipulated that QML cannot, and not for any genuine advantage PWL may
possess.

Of course, if this was not really meant, and it is accepted that predicate calculus
may employ formalisations of any predicate, including such as “Wx”, then the second
problematic feature of such a claim becomes relevant: what, if anything, is a “relational
structure”, and what is it for objects to differ with respect to it? Is a relational structure a
set-theoretic entity of some sort? Is it a structural universal? Is it a complex including
relations and objects (objects such as, e.g. the parts of the thing that has the “relational
structure”)? Is it a linguistic entity essentially involving predicates (as one might suspect
from Hazen’s treatment)? An illuminating translation is difficult when the original subject
matter is obscure. An illuminating translation would also be a trap — if the translation
explains in term of other things what a relational structure is, (and thus, for a start, is
committed to some theory or other about what exactly relational structures are), then it
would contain much more information than the original claim — and thus might not be an
accurate translation of the original claim at all, but the translation of some other more
specific, more metaphysically controversial claim. However, if an unilluminating, but
accurate, claim is wanted, then it is simple enough to produce:

(Ux)(Uy)((Wx & Wy & (Pz)(RSz & Hxz & Hyz)) ⇒ x=y)

(RSx = “x is a relational structure”, Hxy = “x has y”, “Wx = “x is a world”)

If it is objected that this is a non-modal translation, then the obvious response is that the
sentence to be formalised was not a modal sentence. If it is objected that one cannot use a
predicate like “Wx” without buying into PWL, then the reply is that predicate calculus
already had the resources to represent any natural language predicate. If it is objected that
the above sentence quantifies over relational structures, then the response is that the
following is true: either (1) relational structures are not the sorts of things to be quantified
over in first order logic, in which case neither QML nor PWL are capable of such a thing, in which case QML fails no more than PWL here, or (2) we can quantify over relational structures in a first-order way, so there is no problem here. If it is objected that “Hxy” is a mysterious predicate, then the response is that this is deliberate — while “having a relational structure” is mysterious, “Hxy” should be too, where y is a relational structure. A not very satisfactory answer perhaps — but the test was unfair to begin with.

4. Two Challenging Hazen Cases

The next two Hazen cases are more difficult to deal with, for they are cases which can be translated into first-order PWL in a very straightforward way but are quite difficult to translate into QML. In fact, my opinion is that the honest translation of them in QML is second-order, and that this means that there is an expressive advantage which first-order PWL has over first-order QML. However, I will argue that this is not the decisive advantage to PWL that it has seemed to be, and that the second-order QML way of dealing with these cases has some intuitive support that PWL lacks, and also that there are natural-language modal claims which are translatable in second-order QML that have no equivalents in first-order PWL. Furthermore, of course, second-order quantification is hardly a modal theoretical resource, and so as far as the question of whether the disputed claims can be translated into a language which has as its only modal resources the standard modal operators, invoking second-order quantification is completely in order.

The first of the two cases is stated in English by Hazen as

“in every world there is some individual which is also in the actual world” (Hazen 1976 p 39),

although as this stands it is an unfair rendering of the problem, for the same reason as above: on the face of it, the above English claim is a non-modal claim about certain theoretical entities (i.e. worlds) — and is thus not something to worry QML. However, English-language modal translations of this and the second of the two claims are possible, which makes them fit to be serious challenges — see below at pp 64 and 67.

PWL has a quite simple translation of this claim:

\((\forall x)(Wx \rightarrow (\exists y)(Iy@ & (\exists z)(Izx & Czy)))\) if we use Lewis’s PWL (Lewis 1968).
But we cannot provide any translation as simple in QML (Hazen 1976, p 39). One strategy that has been suggested by Hazen himself (Hazen 1976, at p 40) is the introduction of the “is actual” predicate — a predicate possessed by all and only those objects in the actual world — and a predicate that an object has necessarily (or at least essentially) if it has it at all. Then a QML translation becomes easy:

\[ L(Px)(Ex \& Ax) \] (Necessarily there exists something which actually exists).

There are two reasons why this is an undesirable solution, however. The first is that an “is actual” predicate looks like another modal predicate — as “actuality” seems to be a modal concept. Having to introduce a new modal operator to deal with a problem case is already to admit that “it cannot all be done with boxes and diamonds” — but this on its own might be a price we might be prepared to pay. The second reason is that the above Hazen case is just the tip of an iceberg — there is a more general statement of it which cannot be resolved by the use of an actuality operator. It is this more general statement which is the second of the two challenging Hazen claims — and this one is by far the most troublesome.

The more general statement is expressed as the following in Hazen’s PWL:

\[ (\exists w')(w)(\exists x)(Ixw \& Ixw') \]

(Which translates as something like: there is a world such that every world has one of the objects in the first world in it too — without this implying that any specific object in the first world exists necessarily).

To try something like the strategy of introducing an “actuality” predicate would require that a new predicate or other sort of operator be introduced for each possible context (i.e. that there be a predicate or operator corresponding to each PWL world). This is in essence what those who introduce an infinite number of modal operators do — their indexed Vlach operators (Forbes 1989) or indexed boxes and diamonds (Peacocke 1978) function in more or less this manner. To enrich QML with the equivalent of one (or two) operators corresponding to each world in PWL seems to give the game away: it looks a lot like smuggling in possible worlds without the ontological honesty which
possible worlds logic appears to carry on its face. If QML is to be vindicated, it must be vindicated in a way other than the actuality predicate/indexing strategy.

What is to be done? A good place to start is to see if there are any English translations of these two puzzle cases. Hazen says that "It is not at all clear to me that [these two cases] can be expressed in ordinary, as opposed to 'philosophers' English' English" (Hazen 1976, p 41). If he means by this that they cannot be translated into English that does not involve mention of worlds, then I disagree, although if all he means is that the translations would be messy and practically incomprehensible to a "lay reader", then he might have a point. However, the following English translations seem to me to do the job. The first has already been translated above (p 63), but a slightly less ambiguous translation might be

Necessarily one of the objects which actually exists exists, but this is not to say any specific object is a necessary existent.

This is messy, clumsy, and perhaps a little confusing, but this is because English is sometimes ill-suited for clearly distinguishing different modal claims. However, once this English translation is examined, it will be noticed that it looks like it involves second-order quantification ("one of the objects" seems to suggest we are talking about a collection or plurality of objects). Either it can be dealt with by talking of the collection, or set, or class (I will use the generic word "collection" in this discussion) of existing objects, and saying of the collection that it is necessary that something exist that is a member of that collection, or it can be seen as saying of the existing objects that necessarily one of them exists (that is, it can be seen as a case of plural quantification over the existing objects). Once this is realised, it becomes possible to produce second order quantified modal logic statements of the above claim, without needing a special "actually" modal operator.

\[(Pc)(Ux)((Ex \iff x \in c) \& L(Py)(Ey \& y \in c))\]

says that there is a collection such that all and only the (actually) existing things are in it, and necessarily one of the collection exists (but this does not imply that any of the collection exist necessarily). This translation assumes membership rigidity for collections, but this is not an unusual assumption, and I imagine a suitable equivalent to membership rigidity can be defined up by those who wish to take membership as non-rigid. A plurally quantified QML translation can also be generated. A good first stab is:

\[37\text{ Though again, this is denied by Forbes in Forbes 1989.}\]
(PX)(Ux)((Ex iff Xx) & L(Pz)(Ez & Xz))38

(there are some things such that something is one of them iff that thing exists, and
necessarily something (that exists) is one of them).

As it stands, though, it may not be satisfactory. The second occurrence of "Xx"
is in the scope of a modal operator, so its extension may well, for all that has been said,
differ from the first. As it stands, "X" might be functioning more as a variable which
picks out the collection than a name of the collection: it is functioning analogously to a
non-rigid designator, whereas for the translation to be successful, it must be rigid. To
put what is happening here in possible worlds talk, for all the above formula says it is
open that which objects "they" are can be different from world to world — whereas just
as a single object must not be identical with anything other than itself at other worlds, so
there must be a rigid understanding of membership of a plurality, according to which
objects must not be identical with anything other than themselves at other worlds. One
way to resolve this is to introduce into plural quantification the equivalent of names in
order to rigidly designate: so just as a distinction can be drawn between the actual
satisfier of the description “inventor of the zipper” and the possible inventors of the
zipper, the distinction which is responsible for the reading of “the inventor of the zipper
might not have invented the zipper” on which that statement is true, a distinction could
now be drawn between the actual people that invented the laser and the possible groups
that invent the laser, in order that “the inventors of the laser might not have been the
inventors of the laser” can have a sense in which it is true, and furthermore a sense
which is reflected in the formal semantics of plural quantification. Such rigidity can be
represented by adding a condition to the predicate being used to help quantify plurally39,
that if an object is in the extension of X in any world, it is in the extension of X in every
world.40 This would be like producing the rigid version of “the inventor of the zipper”
by producing a predicate (Z) such that someone’s possessing Z is materially implied by
their being the inventor of the zipper, and being in the extension of Z is had by an object
essentially if at all. The amended plurally quantified statement will now go as follows:

(PX)(Ux)(((Ex iff Xx) & (MXx ⊃ LXx)) & L(Pz)(Ez & Xz))

38 I use here the style of plural quantification adopted in Boolos 1984.

39 Alternatively, one could stipulate that the plural variables are always to be interpreted rigidly, and
introduce some other device for modelling non-rigid talk of collections or pluralities. I will not adopt this
style of doing things however, as I think the non-rigid reading is probably the one which accords closer
with natural language.

40 I owe this point to Greg Restall.
This new version ensures that the extension of $X$ will remain the same even in modal contexts. Of course, in putting this extra condition on $X$ we admit that $X$ is not representing a “genuine” property (as it has non-existent objects in its extension in other worlds). This point is also true of the collection-membership relation, above, unless we want to say that a collection can exist at a world where most of its members do not. Even if we do want to say this of collections (as many people have wanted to with sets), there is still a reason for taking the membership relation for the collection not to be “genuine”. For the collection translation to be satisfactory, it must be assumed that the collection cannot have members at any other world that it does not have at the actual world. This was being assumed above, but that assumption becomes questionable if the collection could have different members, as it must have in worlds where not all of the actual members exist, if “membership” is to be genuine, i.e. it is only to hold between existing objects. It is much simpler if it is allowed that it is necessary if true that the collection has a given object as a member. On pain of the necessary existence of both the members and the collection, we must conclude that collection membership in general is not a genuine relation. So both the collection-translation and the plurally quantified translation are committed to an object’s being able to be one of a collection or one of some objects respectively, without existing.

I see nothing particularly objectionable about this however. Surely if there are non-existent objects, we can talk about them plurally as well as singularly. And there is no real problem with considering collections or sets of them either: if there were, then the metatheory of most quantified modal logics would already be in serious trouble (at least that metatheory as used by non-realists about possibilia would be). In any case, it seems plausible to me that talk about sets or collections might be a specific device for plural quantification, and if so, the harmlessness of the latter can be assumed to infect the former.

This second order translation scheme can be used to translate the equivalent of the second of the Hazen cases as well. Remember that in PWL the second case was presented as

$$(\exists w')(w)(\exists x) (Ixw & Ixw')$$

(there is a world such that every world has an object from the first world in it).

41 Here is not the place to attempt to define what a “genuine” property might be — suffice it to say that I suspect that most non-logical run-of-the-mill extensional properties are “genuine”, and that no non-existing object has any of these.
A translation into fairly natural language of this claim is that there is a collection of copossible objects such that necessarily one of the members of that collection exists (or alternatively, there are some copossible objects such that necessarily one of them exist) — and furthermore that the collection is such that it is possible that its members be all of the objects (as possible worlds are maximal, after all). Alternatively, if it is thought that the “there is” and “there are” phrases in the English sentences imply ontological commitment (or imply that such objects are actual), then the translations can be modified to say that either

i) There could be a collection of objects, such that the members of the collection were all the existing objects, such that necessarily one of members of that collection exists, or

ii) There could be some (co-existing) objects such that they could be all of the existing objects and necessarily one of them exists,

depending on whether talk of collections or plural quantification is to be preferred.

In either translation of course, the claim made is not taken to imply that any member of the collection (or any one of the objects) in particular necessarily exists. These modalised claims are also likely to be more faithful to the original PWL version, as an existential quantification over worlds at the start of a formula is most naturally read in “modal talk” as “it is possible that”. Let us therefore use (i) and (ii) as the claims to be formalised.

The formalisation of both is rather easy. The first is

\[ M(Pc)(Ux)(x \in c \iff Ex) & L(Py)(E y & y \in c) \]

or perhaps

\[ (Pc)M(Ux)(x \in c \iff Ex) & L(Py)(E y & y \in c) \]

and the second is

\[ M(Px)(Ux)(((Xx \iff Ex) & (MXx \Rightarrow LXx)) & L(Pz)(Ez & Xz)) \]

or perhaps

\[ (Px)M(Ux)(((Xx \iff Ex) & (MXx \Rightarrow LXx)) & L(Pz)(Ez & Xz)) \]

(There is really no difference between the two versions of each, when it is remembered that the neutral logic being used validates the Barcan formulae). The translations into collection-quantified and plurally quantified modal logic are both very straightforward — they are more or less just the possibilised versions of the first Hazen case, as one might expect them to be.
5. **Defending the Use of Second-Order QML**

The crucial question governing the acceptability of these translations as genuine rivals to PWL, apart from the question of what modal resources are employed, is the following: does the fact that these translations have to be second order mean that PWL does have a genuine advantage over its QML rival? After all, was it not the case that the original challenge was made to first order quantified modal logic? Might it not be charged that these natural language and second-order quantified statements, however interesting they may be, actually vindicate the claim of PWL to be superior in virtue of its greater expressive power, not count against it?

Of course, it is undeniable that PWL has the advantage (if it is an advantage) of being able to translate in first order terms claims that require 2nd order locutions in the QML I am using. I will argue that this does not count in PWL’s favour, however, for two reasons. The first is that the normal modal English translations of these two Hazen claims are apparently second-order, or have natural second-order renderings — so a second order formalisation of these claims is more faithful to our natural language than a first order language which achieves the result by introducing objects (worlds) and predicates (“is a world”, “is in” etc.) which are alien to the natural language. The second is that the Hazen cases can be seen in the context of a more general category of modal claims, which are made using second-order modal locutions in English, and not only have no first-order QML translations, but have no first-order translations in PWL either. I will argue that an account of this general class of claims will reveal that the second order formalising strategy must be the correct general strategy for translating these cases, and show how this account, once provided, shows the Hazen cases to be special cases of this more general category, and that their ability to be translated in first order PWL is more or less coincidental. I will thus argue for the expressive incompleteness of (first-order) PWL with respect to second order QML when it comes to ordinary language modal claims, and furthermore demonstrate why the broader understanding of these cases makes the second order QML treatment of the Hazen cases is better motivated than the PWL version.

Recall that the English equivalents of the Hazen cases were claims made about the modal properties of certain collections of objects or of the modal properties of pluralities of objects. In particular, recall alternative formulations of the English equivalent of Hazen’s second, more general case:

i) There could be a collection of co-existing objects such that they could be all the objects and necessarily one of members of that collection exists, or

ii) There could be some co-existing objects such that they could be all the objects and necessarily one of them exists.
There are other claims to do with the modal properties of certain collections of objects (or pluralities of objects) which look to be of the same kind as the above claims. One example is the following:

(1) There are (at least) two collections of objects such that necessarily none of the members of one can co-exist with all of the members of the other, and vice versa.

This claim as it stands is ambiguous, but I intend the example to be the claim expressed by the above which is not the claim that none of one can coexist with any of the other. The plurally quantified version of this claim can be made to sound more like ordinary language:

(1') There are some objects, and some other objects, and necessarily, if all of the first exist none of the second do, and vice versa.

The formalisations of these claims are fairly straightforward:

Formalised as a claim about collections:

\[(Pc)(Pc')\land(((Uw)(w\in c \supset Ew) \supset (Ux)(x\in c' \supset \neg Ex)) \land ((Uy)(y\in c' \supset Ey) \supset (Uz)(z\in c' \supset \neg Ez)))\]

and as a claim about pluralities:

\[(PX)(PY)(Uv)(((MXv \supset LXv) \land (MYv \supset LYv)) \land ((Uw)(Xw \supset Ew) \supset (Ux)(Xx \supset \neg Ex)) \land ((Uy)(Yy \supset Ey) \supset (Uz)(Xz \supset \neg Ez)))\]

Another example of a claim to do with modal properties of collections or pluralities of objects is the following:

(2) There is a collection of objects (or: there are some objects), each of which is co-possible with at least one of the others, and necessarily one of them exists.

This is translated as

\[(Pc)(Ux)((x\in c \supset ((Py)(y\in c \land \neg(x=y) \land M(Ex \land Ey)))) \land L(Pz)(Ez \land z\in c), \text{ or}\]

\[(PX)(Ux)((Xx \supset ((Py)(Xy \land \neg(x=y) \land M(Ex \land Ey)))) \land (MXx \supset LXx)) \land L(Pz)(Ez \land Xz),\]

depending on whether we prefer to talk of collections or plurally quantify.
Nevertheless, we do still run into these problems at some point. Lewis does not believe in the existence of impossible worlds, or the existence of most impossibilia.\textsuperscript{32} The round square cupola on Berkeley college, for example, is not one of the things that exists, and neither is the highest prime number. Once we get to this point, Lewis’s PWL does no better than classical logic to translate these sentences. For instance, call the highest prime number “$h$”\textsuperscript{33}. How are we to translate the claim “$h$ does not exist”? We can try $(\forall x)\neg(x=h)$ — but that will come out as false (or, on some interpretations, meaningless because “$h$” will not have been assigned an extension), whereas we want it to be true that $h$ does not exist. We could use our PWL’s extra resources and try

$$(\forall x)(\forall y)((Wx \& Iyx) \supset \neg(y=h))$$

and this appears adequate: it is not ruled out by the postulates, and says that $h$ does not exist in any possible worlds, which might be good enough. However, when we take into account that Lewis’s counterpart theory, at least, “had best be understood as quantifying over only possible individuals” (Lewis 1983), then there can be the problem that the model theory will rule out the above sentence automatically, if the domain is restricted to worlds and objects in worlds. It seems to be an unfortunate gap in Lewis’s logic that the intended model theory counts as theorems some things which are not derivable from the proof theory\textsuperscript{34}: but once the proof theory is rectified to reflect the intended model-theoretic interpretation, it turns out that “$(\forall x)(\forall y)((Wx \& Iyx) \supset \neg(y=h))$” is false, at least when $h$ is not a world. Again, this indicates that the translation is inadequate — for surely the claim that $h$ does not exist is true, where $h$ is meant to be the highest prime number.

Classical logic cannot translate the first three cases because they all involve talking about what is not: the first, to say of something that it is not the case that it could fail to exist, as opposed to other things which could fail to exist. Since classical quantified

\textsuperscript{32} On the other hand, there are some impossibilia that he does admit the existence of: mereological aggregates of parts of different worlds (Lewis 1983 p 39-40). These are very much a special case when it comes to impossibilia, of course.

\textsuperscript{33} This example will seem infelicitous to those who believe numbers are not fit for objectual quantification (perhaps they are to be the subject of 2nd order quantifiers only, or some such). For those people, feel free to substitute some other impossible description as the definition of “$h$” (A round square cupola, or the cricket ball which is both red and not-red all over, or some other such thing). I am also assuming here that names are not disguised definite descriptions, or some such.

\textsuperscript{34} For instance, if the objects in the models are restricted to worlds and objects in worlds (that is, it is restricted to possible objects), then no model will allow that there is a (non-world) that is not in any world. Yet it can be shown that

$$(\exists x)(\neg Wx \& (y)(z)((Wy \& Izy) \supset \neg(x=z)))$$
does not follow from Lewis’s postulates in (Lewis 1968).
modal logic has no way of saying truly that anything does not exist (or even possibly
does not exist), it is no surprise that it cannot make any contrast between saying of
something that it necessarily exists, and saying of something merely that it exists. This is
why (\exists x)L(\forall y)(x=y) seems to mean both that something exists necessarily, and just that
something exists — notice it is logically equivalent to, and derivable from, the claim that
(\exists x)(x=x). At least this is so in the modal system Hazen offers as the "rival" to his PWL

PWL runs into problems with existence for just the same reasons that classical
logic does, and only manages to translate some claims classical logic cannot because it
takes so many more things to exist (whereas classical modal logic, in effect, can only say
that the actual objects exist). These cases, then, should not be seen as a particularly
strong reason to embrace PWL instead of QML, especially since making QML "free"
gives it the resources to translate not only the first three Hazen formulas about identity,
but can translate the claims about existence that PWL cannot either. For the free QML I
have suggested, "h does not exist", receives the perfectly simple

\[ \neg E_h \]

as its translation, and if we wish to say not only that h does not exist, but could not have,
then

\[ L\neg E_h \]
captures this quite adequately too. And neither \( \neg E_h \) nor \( L\neg E_h \) are trivially false according
to this neutral QML. Neither are claims like \( L(\forall x)(\neg E x) \) or \( (\forall x)L(\neg E x) \) trivially true
either, whereas it might be thought that it is trivially true that something (e.g. the round
square) does not exist. This might seem to be a lingering problem. If it is, then it is
certainly a further problem which would arise a fortiori for a PWL — but maybe it is not
a great problem. It has often been thought that existential commitments were not a matter
of logic — maybe commitments to objects' non-existence should not be either.

This advantage which neutral QML has over PWL is one that might not be
accepted by some defenders of PWL, even if they are prepared to quantify over what is
not. Some philosophers feel that there are especial problems with quantifying over
impossibilia which makes it impossible or nonsensical to quantify over such entities,
even if quantification over possibilia without ontological commitment is acceptable.
People who have held this view apparently include N. B. Cocchiarella\(^35\), and the Routley
of 1966\(^36\). Even if this should turn out to be the case, it is still true that a neutral QML

\(^35\) discussed in Bencivenga 1986, at pp 391-392

\(^36\) Routley 1966, pp 259-260, although even here he says that quantification of impossibilia is
unacceptable only given certain fairly classical assumptions, and that, for example, there are Meinongian
systems of quantification which quantify over impossibilia without falling into inconsistency or
incoherence. By Routley 1980, of course, Routley had adopted Meinongian quantification and defended
that quantifies only over possible objects can match the expressive completeness of PWL in these three cases.

Strictly speaking, to show that QML can match the resources of PWL it is not even required that we be allowed to quantify neutrally in order to match the expressiveness of PWL, at least if modal realism is true. For then, provided one interprets the predicate “E” as meaning “actually exists” (understood de dicto rather than de re, so that the predicate is equivalent to the predicate “exists in this space-time” where “this” is understood non-rigidly) and restrict the domain of quantification to the realm of all existents, actual and merely possible, then the QML translations offered above show that a QML matches the expressibility of PWL, at least in the cases of the three cases offered by Hazen. However, while this might be the case as a point of logic, it is not necessarily going to be interesting metaphysically: as requiring the resources of modal realism (or some other theory which postulates the existence of possible worlds and merely possible objects) to justify one’s QML is just as good an argument for the existence of possible worlds as the argument from showing that it is required that one use the resources of a theory committed to the existence of possible worlds in order to justify one’s PWL to the conclusion that possible worlds exist. However, the observation that QML is equivalent in expressive power to PWL when it comes to formalising natural language modal claims does speak to the semantic issue of whether modal operator talk or quantification over worlds talk is semantically prior: and so it may still be a useful observation when it comes to doing the semantics of natural language modal discourse. In future, then, it can be understood that neutral phrasing of various claims is not strictly necessary to establish that QML has the same expressive power as PWL if it is accepted that possible worlds and possibilia all exist, even though I will not bother to stop and make this point explicitly again, because as a matter of fact I doubt the existence of mere possibilia, and am not even convinced of the existence of non-actual possible worlds.

Note finally that if I am right about the uses of the devices of neutral quantification and the existence predicate, then are useful for non-modal theoretical tasks as well. Even leaving aside hyperintensional uses (such as solving the Frege-Geach problem, accounting for cross-references in radically false theories and so on) — which I will not dwell on because their status is a vexed question and the issue of whether or not the hyperintensional is to be analysed in modal terms is also up in the air — there are several non-modal uses of these devices. One is to say that some things do not exist: existence does not seem to be a modal matter, and neither does its absence. And being fit to be neutrally quantified over is plausibly not a modal matter, since the possible and the impossible are equally to be found in the quantifiers’ domain. Whether one wishes to perform such tasks with one’s theory is another matter: but even those who are hostile to quantification over impossibilialia (see pp 83-84).
these tasks or thinks them confused should be able to agree that they are not \textit{modal} theoretical tasks.

Now that the problems of existence have been addressed, let us move on to the more difficult of Hazen's cases. Hazen in fact grants that certain first-order systems can express claims of the sort I have been dealing with so far — the logic including "outer quantifiers" (which is a version of a neutral logic, albeit one restricted to quantification over possibilia) is the one he considers. He even notes that the "outer quantifier" logic he considers is capable of non-trivially expressing statements which cannot be other than logically false according to his PWL, including the very claim that I was discussing above — that a free logic can express the claim that something necessarily does not exist. However, he takes the next three series of examples to provide modal statements which he can translate but QML, even a free QML, cannot. The first is a case I will argue is unfair, but the next two are genuinely problematic.

3. Hazen Case Number Four: A Case To Be Set Aside

The fourth case Hazen raises is the claim that "sameness of relational structure is a sufficient condition for the identity of possible worlds". I take it to be an unfair example of the expressive advantage of PWL over QML for two reasons. The first is that he cannot offer a translation of this even into his PWL. The best he does is offers us a translation scheme, and even it only works, according to him, if the number of "non-logical" predicates is finite. And even this claim about his translation scheme is false, once it is realised that it is not the number of predicates that is at issue, but the number of properties and relations that are in existence (or could be in existence) — and anyone but a predicate nominalist should concede that these two numbers do or at least can diverge. To be fair, Hazen claims only that the scheme he offers "goes some way" to translating the above claim — but if, as it seems, it cannot go all the way, then it is unfair to expect QML to do so.

As if the untranslatability of the claim into PWL was not bad enough, there is a second reason to think that the above claim is not a suitable candidate for showing PWL's expressive superiority with respect to modal claims: the claim is already phrased as a non-modal claim in a language quantifying over worlds — it is not a modal claim in English, nor that it is clear that it can be recast as a \textit{modal} claim in a natural language. It is therefore not to the discredit of a logic set up to formalise modal claims of our natural language that it fails to provide a formal (modal) translation for this claim. To be fair, the task Hazen sets himself in (Hazen 1976) is to show that there are some claims formalisable in PWL that are not formalisable in QML — not that these claims are modal in character, or that they are natural language modal claims — so this case may not be "unfair" for this second reason given his project, even though it is if it is viewed in terms of an argument to the effect that PWL is to be preferred because it captures natural-
3. The Second Problem: The Incompleteness of the Modal Fiction

The problem of the incompleteness of the Modal Fiction is the first of the two ontological worries I will discuss, and is similar to an objection Lewis has brought against the position he labels sparse linguistic ersatzism (Lewis 1986c, pp 142-165). The point comes when we consider what it is for something to be "true according to the fiction". Maybe this is a primitive for the modal fictionalist. However, "primitive" does not mean the same as "totally mysterious". At least a little can be said even about primitive notions.

One thing that seems clear is that a sentence, or proposition, or whatever is true according to a fiction if (but not necessarily only if) that sentence or proposition is part of the fiction, in the sense that it is one of the assertions of the fiction. I will call the things true in a fiction in virtue of their being one of the assertions which make up the fiction the explicit content of the fiction. (This usage is different from the way the term "explicit content" is used normally in discussion of fiction (e.g. in Lewis 1978, p 41) but is rather similar to Lewis's notion of "explicit representation" which he employs in On The Plurality of Worlds in his discussion of ersatz theories (see Lewis 1986c, pp 150-151).

Now, it is clear that merely the explicit content of the candidate fictions will not be sufficient. If anything like the range of modal claims which we take to be true are actually true, we will require the content of the fiction to be at least infinite. There are an infinite number of possibilities, and for a modal fictionalist this means that there an infinite number of statements which are true according to the fiction. Now, take a look at the fictions actually put forward as candidates for being the Modal fictions. Rosen suggests a modified version of Lewis's On the Plurality of Worlds, (in Rosen 1990), and Armstrong in his (Armstrong 1989, at p 50) claims that there is a "great fiction" which says that there are many other fictions that each describe a possible world. This "great fiction" can be identified with what he says about the arrangement of possible worlds in A Combinatorial Theory of Possibility. The other fictions supposedly described by the great fiction need not concern us as much, since according to Armstrong they do not exist, but are themselves only fictional entities. Both On the Plurality of Worlds and the fictional utterances in A Combinatorial Theory of Possibility are fictions that can be read in a day, and neither's explicit content is detailed enough to support the possibility claims we want to make. Furthermore, no fiction which humans could ever produce could have enough explicit content to support all the modal claims which we suppose to be true — annoying things like finite matter and energy will get in the way of such a project.

Note that the Incompleteness Objection I will be discussing is quite different from the "incompleteness problem" discussed by Rosen in (Rosen 1990, p 341). Rosen is concerned with deciding issues about which the Modal Fiction is silent but about which there are determinate facts of the matter according to modal realism, whereas my objection concerns difficulties involved with modal fiction being supposed to make many of the claims about which it is not meant to be silent.
So there is not enough of the Modal fiction for it to explicitly make all the assertions we need. What we do have in these modal fictions, however, are a lot of universally quantified sentences (such as those found in principles of combination or other statements which apply to “all worlds”). “Aha”, the fictionalist can say, “these sentences imply all the content we need”. This is the sort of thing we can usually say about fictions — there is probably no sentence in Tolkien’s writings which states “Merry was a hobbit and Shelob was a spider”, but this conjunction is implied by his fiction, as it says in one place that Merry was a hobbit, and in another that Shelob was a spider (in fact, these claims may in turn have only been implied, as it is unlikely that this information was conveyed in such a bald fashion). Here is how finite works can have true according to them an infinite number of propositions — they can imply them all!

This is not really open to the strong modal fictionalist, however — implication, a modal notion (some might argue the modal notion), is meant to be explained in terms of what is true in the story about possible worlds, not the other way around. If we are going to embrace implication as something relatively more fundamental than the Modal Fiction, or “relatively primitive” as compared with it (if relative primitiveness makes sense), then why not just be a timid fictionalist, especially since it looks like other modal notions are definable in terms of implication (e.g. \( P \) is possible iff it is not the case that \( P \) implies \( \neg P \))? If, on the other hand, the modal fictionalist cannot appeal to the implications of the fiction, it seems that the content must be explicit — in which case the fictions offered are woefully inadequate. To state the problem succinctly: the fictions do not contain sufficient explicit content to perform the function required by strong modal fictionalism and the theory cannot rely on any implicit content they might be supposed to have. In Rosen’s case the worlds are not described in enough explicit detail, and in Armstrong’s case it is the descriptions of the worlds which are not described in enough explicit detail.

The timid modal fictionalist, on the other hand, can rely on implicit representation and implication in producing their theory of the modal fiction. Since the timid fictionalist does not purport to be offering any sort of reductionist analysis of modality, there is no threat involved in relying on modal notions to characterise the modal fiction, as there is for the strong modal fictionalist. What to say about the broad modal fictionalist is more difficult. In one sense the broad fictionalist can also without circularity avail themselves of modal idiom in characterising the fiction of possible worlds, since that fiction is not intended to provide a reductive explanation of modality for the broad fictionalist either. However, the broad modal fictionalist will encounter the same problem when it comes to

48 Thus I think that the “syntactic/inferential” account of the modal fiction (and any account relevantly similar) is in even worse shape than Lycan (Lycan 1993, p10) argues, since Lycan admits that an adequate fiction might be able to be provided through using universal quantification over “all combinations” and such like.
their fiction of the truth of modal claims themselves. That fiction will need to explicitly 
represent its content, on pain of vicious circularity, for the same reason as the strong 
modal fictionalist’s story about possible worlds must. Since there will probably be an 
infinite number of modal statements which are true according to an adequate broad modal 
fictionalist’s fiction about modality, this is a tall order for the broad modal fictionalist. So 
the Incompleteness problem, in a slightly different guise, is a problem for broad modal 
fictionalism as well.

4. The Third Problem: Propositions

The third problem for the Modal Fiction is also an ontological worry of a different sort. 
It arises when we come to consider what propositions are: in other words, what are the 
things that are true in the fiction supposed to be? It seems that at least two strong modal 
fictionalist positions come to grief when faced with this question: Rosen’s and 
Armstrong’s. As always, there are repairs that might be suggested, but once again, these 
come with a price in terms of plausibility for the fictionalist theory.

The modal fictionalism in Gideon Rosen’s “Modal Fictionalism” is the one most 
obviously in trouble. The strategy there is to take advantage of the benefits of David 
Lewis’s theory of possible worlds, while avoiding paying the ontological cost, by being 
fictionalist rather than realist (Rosen 1990, p 330; Rosen 1993, p 73). Now, one of the 
benefits of Lewis’s system is that he gives us an account of propositions: propositions, 
for Lewis, are sets of possible worlds. Since possible worlds do not literally exist for the 
fictionalist, propositions (being sets of possible worlds) will not exist either — either 
that, or each will be either the null set or the unit-set of the actual world, which does not 
nearly individuate them adequately. If the Fictionalist is left without propositions, there is 
nothing to be true according to the fiction, and thus there will be no modal truths — a 
clearly unacceptable conclusion for the Rosen-style strong fictionalist.

Of course, these propositions may well exist according to the fiction, but to try to 
use this to get around the problem leads to a vicious regress. Because then it is not true 
simply that a proposition \( P \) is true according to the fiction, but only that “according to 
the fiction, \( P \) (exists and) is true” is true according to the fiction. And even that will not 
be true simpliciter, for the proposition “according to the fiction, \( P \) exists and is true” does 
not actually exist, but only exists according to the fiction, so all we can say is that 
“according to the fiction, the proposition “\( P \) exists and is true according to the fiction” is 
true according to the fiction”’’ is true according to the fiction. And we cannot stop here 
either — we must again qualify it. The problem is that we never arrive at a true statement: 
only statements which must be further qualified. This means that there is no acceptable 
paraphrase or expansion of “\( P \) is true according to the fiction” which is actually true — so 
we still come to the conclusion that nothing is true according to the fiction. That the
propositions supposedly exist according to the fiction is not something that will save the fictionalist who wants to embrace the Lewis story as their fiction.

Armstrong, too, will have problems when it comes to the question of what the propositions involved with the Modal Fiction are. Perhaps “propositions” is not quite the right word for the things which make up the fiction according to Armstrong, as Armstrong rejects ontological commitment to propositions (see Armstrong 1973, p 46). Let us call the things which appear in fictions “assertions” instead, as these are a species of entity Armstrong does allow into his ontology. An account of fictions as sets or collections or other organisations of assertions will cause the following difficulty for Armstrong: Armstrong admits, as anyone who is not anti-realist in the broad sense about modality, that there are possible states of affairs that have never been mentioned or described in thought or in speech and in all probability never will be. Call one of these US (for unmentioned state of affairs). Yet for this to be so, then there must be an assertion which is true in the fiction to the effect that there is a little fiction which contains US. Now, no actual assertions exist which assert that US holds, so there is nothing actual to be the part of the fiction which asserts US. On the not unreasonable assumption that if US has never been mentioned, neither has any assertion which means “there is a little fiction that has true according to it US” been uttered, it follows that there is no assertion of the appropriate sort corresponding to the possibility of US: so it seems that US is not possible after all. And again, an appeal to the fictional existence of mentions of US will not help, for the same reason that appeal to fictional possible worlds as the members of propositions did not help the Rosen-style fictionalist. The problem for both Rosen and Armstrong is that some propositions (or assertions) which they need to be true in their fictions in order for various states of affairs to be possible do not even exist, and so cannot be true according to their modal fictions.

This net can be cast even wider still: any account of propositions which unavoidably involves modal notions, for instance, ones that make reference to non-actual acts of meaning something, or classes of inscriptions or utterances that include mere possibilia, or talk of “ideal” languages or descriptions of the world, where this ideality is cast as some sort of modal notion, and so on, will be incompatible with strong fictionalism.

This objection is not meant to be a decisive one: it only works if the modal fictionalist is guilty of smuggling something like modality into his or her account of propositions. However, it seems clear that we cannot just restrict ourselves to propositions corresponding to statements which have actually been made (or, more generously, which will be made at some time). For there are likely to be very few statements of the form “According to the Modal Fiction, P is true at some possible world” — and at the price of seriously restricting the possible (or being anti-realistic about possibility in the way I considered above), we want our theory to say that possibly P is
true even when no-one can be bothered to utter the relevant statement about the fiction, or fictively utter the relevant statement in the context of the fiction. What one can do to avoid this objection, of course, is say that all possible propositions, both expressed and unexpressed, are actual. This Platonism about propositions will likely take propositions to be abstract objects of some sort: maybe sets of some special sort, maybe “meaning stuff” in Plato’s heaven — who knows? I fully admit that this sort of theory of propositions avoids the objection.

In fact, if one is to be Platonist about propositions, one can to some extent avoid all the problems raised in this chapter. The Modal Fiction, if it is a collection of propositions or some set or ordered set of propositions, will not then be artificial — no matter what humans do, or even if we do not exist at all, still the contents of Platonic heaven will exist. The artificiality objection would then have to be recast: the artificiality objection might be that it is somehow up to us which of the innumerable fictions about pluralities of worlds comes to be the fiction that is connected in the appropriate ways to our modal claims. But this objection would not be quite as strong, as it is open to the fictionalist to claim that it is an objective matter which fiction propositions about possibility and necessity are about (or are made true or false by), and that this is not the sort of thing that is affected by the musings of philosophers, or, indeed, the existence of sentence or life at all (though an explanation which does not rely on any features of us which explains why this particular abstract fiction is the modal fiction is still a challenge). Furthermore, the fictionalist who is a Platonist about propositions can supply the missing ontology to avoid the objection based on the supposed incomplete content of the modal fiction: for Platonic sets of sentences can be infinite, and a Platonic fiction can have all of its assertions being explicit, and indeed describe every one of the infinite possible worlds in total detail without having to rely on implication, unlike any story which could be written into a book by mere inhabitants of the world of change and decay. Of course, the problem of how one huge and infinite collection of propositions came to be the Modal Fiction rather than any other will remain — and it will be no good for the strong modal fictionalist to specify that it is the collection which explicitly represents the entailments of something they say, (or indeed give any other modal specification of the collection of Platonic propositions) on pain of vicious circularity.

Why should a strong modal fictionalist not take this approach to propositions then? The first point is that commitment to such Platonism about propositions might be thought to be an ontological price too high to be worth paying. It is difficult for me to see what sorts of things actual abstract propositions might be, and if a reduction of such objects to more orthodox entities, such as sets or facts or universals cannot be carried out, it seems we have, through our analysis, just swapped primitive modality for equally mysterious primitive propositions — and if this is so, it is not clear that we have a net saving in the parsimony of our total theory. Thus even if the question of propositions is
answered in a non-modal way, the cost of providing such an explanation of propositions may itself count as an argument against strong modal fictionalism. The second consideration is that this brand of strong modal fictionalism probably does not even deserve the name. It differs only slightly from representational abstractionism — the only difference is that a representational abstractionist will identify worlds with what the fictionalist takes to be descriptions or representations of (admittedly non-existent) worlds. A strong fictionalism which is Platonist about propositions will probably suffer in comparison with such abstractionism in fact, for it will presumably face any problems which bedevil such abstractionisms, and will have the additional problem of needing to invoke the troubling notion of truth-in-fiction.49 This second point is not an objection to dealing with modality with an ontology of Platonic propositions *per se*, of course: the point is rather that one who employs this approach is an abstractionist rather than a strong modal fictionalist, or at worst a fictionalist who hopes to take advantage of Platonic propositions holds a close relative to representational abstractionism which shares all of its costs and has an additional cost of its own as well. So while Platonism about propositions can be invoked by a strong modal fictionalist to avoid the three problems discussed in this chapter, it would seem to undercut the original motives for modal fictionalism: neither the desire to avoid ontological cost nor the desire to develop an alternative to ersatzism are well met by this strategy.50

The strong modal fictionalist can take comfort from the knowledge that s/he is not alone in finding the objection based on propositions a problem. It is not entirely obvious what a timid modal fictionalist could say if they wanted to avoid unreduced Platonic propositions — and the timid modal fictionalist clearly cannot rely on the propositions-as-sets-of-possible-worlds account on pain of circularity. However, the timid modal fictionalist is in a better position insofar as s/he is able to employ modal resources in constructing a solution to the problem. I will not be attempting to do this job for the timid modal fictionalist here, but I believe the timid modal fictionalist has a better chance of solving this problem than the strong modal fictionalist who must restrict herself to non-modal resources.

49 Of course, when it comes to comparing total theories the abstractionist will need to give an account of truth-in-fiction somewhere along the line: but the abstractionist will then have the advantage of being able to use modal resources or talk of possible worlds to help non-circularly explain the notion of fiction (s/he or he could, for instance, simply take over the account of fiction given in Lewis 1978).

50 Points quite similar to this are made against the Platonism-about-fictions strategy in (Lycan 1993, footnote 18 on p16)
5. An Assumption Relied on By the Incompleteness Objections, and a Defence Thereof

Before I leave the topic of ontological incompleteness, I should mention an assumption that runs through both versions of the incompleteness objection and which both objections rely on to some extent, and mention why I think the fictionalist reply to these objections which involves the denial of the assumption is not plausible.

The assumption is that when we say that a certain proposition is true according to a certain fiction, what we are saying is that a relation of a certain sort holds between a fiction and a proposition. It has been suggested to me that instead of seeing it as asserting that the relation of "being true according to" holds between a proposition and the Modal Fiction, rather it is the application of an irreducible monadic predicate to an object. This strategy, in different guises, supposedly provides an answer to at least one of the versions of the incompleteness objection, and perhaps to both. For if "is true according to the Modal Fiction" is an operator on propositions that does not imply any sort of relation (such as the part-of relation, or an implication relation) between a proposition and the Modal Fiction, then it does not matter if the fiction does not have such propositions in it or even implied by it. By the same token, if "P is true according to" is a monadic predicate applied to fictions that does not require that there be any such proposition as P, then the non-existence of such propositions is no bar to saying, for some P, "P is true according to the modal fiction". Obviously, this strategy cannot be used simultaneously to answer both halves of the incompleteness problem. Apart from the fact that this strategy just screams "ad hoc", and flies in the face of a plausible first stab at analysis of "truth in fiction", it faces another objection: either of these irreducible monadic predicates are utterly mysterious, and an analysis of an area of discourse into an explanation in these terms lacks genuine explanatory power. If these predicates are the best things strong modal fictionalism can come up with to analyse modal discourse, then it seems to me that it would be better to embrace sui generis modality rather than make our theory more complex just to introduce such sui generis mystery further on.

6. A Final Problem: Redundancy

The final problem for Modal Fictionalism is possibly the least decisive, but in terms of motivating an abandonment of strong modal fictionalism it may be the strongest. Recall that when I was setting out the artificiality objection I noted that the modal fictionalist was not laseiz faire about the content of the modal fiction, but rather this content was constrained by various facts about actuality: though whether these things be internal connections between universals and facts about the nature of states of affairs, as Armstrong might take them to be, or facts about our imaginative practices combined with
an encyclopedic listing of actual truths, as Rosen suggests the actual grounds might be, or whether the features of this world which justify the use of one fiction about worlds rather than another are something else altogether, will vary between different versions of strong modal fictionalism. Call these truths about the world, just for the sake of this discussion, the "basic truths". The objection then is this: given that there are truths about the world (of some sort or another) which the fictionalist admits ultimately determine which modal claims are true or false anyway, why invoke the fiction as well — why not simply say that it is the basic truths that are the truth-makers for the modal claims, or that modal claims reduce to claims about such basic truths, or in some other way the basic truths are the explanations of the modal truths? Why not leave discussion of the fiction out altogether? Why is the detour through talk about fictions and truth in the modal fiction not a redundant detour?

It must be admitted that a theory that only needs to mention the basic facts and the modal truths is simpler than one that needs to talk about the basic facts, modal truths and as well discuss the rather mysterious and problematic entity which is the Modal Fiction. The onus is on those who would add this extra factor into the analysis of modality to show why it is needed, especially in the light of the fact that it raises many new difficulties, including the ones previously mentioned in this chapter. This then, might not be so much of an objection as a challenge: while I am not arguing that a detour through the fiction in order to come to the modal truths could not be unavoidable, I charge that it has not been shown that it is in fact required.

7. Conclusion

As I said near the start of the chapter, this discussion of problems was by no means meant to be exhaustive: but it is my opinion that these three are enough on their own to make strong modal fictionalism very unattractive. Strong modal fictionalism introduces artificiality where we would wish to avoid it and it raises ontological problems, both through the inadequacy of what limited things like fictions can represent and through forcing us to give a non-modal account of propositions. As in virtually any area of philosophy, salvages are possible: but in both the cases of Platonism about propositions and introduction of sui generis complex "according to" operators the salvages themselves are very unattractive. Fictionalism about possible worlds could turn out to be a very useful approach, since it is undisputable that we sometimes talk as if there are such things, and yet it would be nice if we could avoid the problems of actually embracing the view that this talk is to be taken at face value. If Fictionalism is to be accepted, however, we must not embrace it as it is often presented: we must not embrace strong fictionalism about possible worlds.
Chapter 5 — Representing Aliens

This chapter concerns itself with a problem about how it is possible to represent things which do not actually exist. As we will soon see, the most difficult aspect of this question boils down to the problem of how it is possible for there to be representations of alien universals — that is, how it is possible to represent properties and relations which do not actually exist, and which are not constructible from those which do, but which nevertheless possibly exist. Two versions of this problem will be discussed — the more general and easier question of how we might be able to produce representations of alien universals at all, and the more specific and more difficult question of how representation of alien universals can be accomplished in the context of so-called “linguistic ersatzism”, a version of representational abstractionism about possible worlds. David Lewis has claimed that there is a serious problem for linguistic ersatzism in this regard (Lewis, 1986c, p 159-160), and though he does not discuss it, the problem arises in a very similar form for modal fictionalist theories. I will show however that this problem can be surmounted in both cases. In so doing, however, I will show that a linguistic ersatzist theory must be formulated somewhat differently from the traditional way. Since both modal fictionalism and linguistic ersatzism are views of possible worlds I feel sympathy for, it will be useful to deal with one of the main objections which they face.

In choosing to discuss this topic, I am not addressing the only interesting problem which comes up for alien universals in the context of modality. Alien universals cause problems for combinatorial accounts of modality, for example the theories in Armstrong 1989 and Skyrms 1981. In part these theories of modality suffer from the problem I will discuss — The theory outlined by Skyrms in Skyrms 1981 can be read as a linguistic ersatzism (though Armstrong 1989 p 46 claims that Skyrms told him that in fact he (Skyrms) holds a modal fictionalist position), and Armstrong’s theory is a modal fictionalism. However, these theories face further problems from alien universals. In Armstrong’s case, it is a problem which flowed from his principles about the truthmakers for modal claims (in particular, he takes the components of the truthmaker for a claim that a given property or relation is instantiated to include at least the basic components of the property or relation itself). Skyrms, on the other hand, claims to be able to deal with

51 This definition of what it is to be an alien property or relation is substantially the same definition as that offered in Armstrong 1989 p 54. This, in turn, is a modification of the definition of what it is to be an alien property provided in Lewis 1983 p 364. Lewis offers a definition of the relation of alien-to-a-world — the current definition is produced by considering as alien simpliciter those properties and relations alien (in Lewis’s sense) to the actual world. Also, Lewis might prefer not to say that universals alien to a world do not exist at that world, but rather are just not instantiated at that world. See the next section for details.
alien universals, but not to be able to treat them as full-bloodedly as the universals instantiated in this world (see Skyrms 1981 p 148). Why he makes this claim is unclear, as is much about his position. I will solve the difficulties facing these positions which follow from their being fictionalist or ersatzist, though the specific problems these theories bring upon themselves will not be dealt with.

Note also that by "universal" I do not mean anything terribly technical in this chapter: my use of "universal" in this chapter is a handy short-hand way of saying "property or relation". As a matter of fact I hold a view of properties and relations which sees them as universals in the more technical sense of the word, but that will make little difference to this chapter. The companion piece of jargon "instantiates" should also be read in a fairly minimalist fashion — an object (or collection of objects) will instantiate a universal when it has that universal as a property (or they stand in the relation which is that universal). I have been and will be talking as if a given universal only exists in the worlds in which it is instantiated — though in the next section I will show that nothing vital to this chapter hangs on this conception of universals either.

1. The Problem

What, then, is the problem I am discussing? The problem only arises if we have a theory of the meaning of nouns which relies on either some causal or direct connection to the things "named"52, or alternatively a kind of general empiricism which requires that things enter into our experiences, directly or indirectly, before we can become cognisant of them or formulate specific truths about them. I will not discuss either approach in detail, but suffice it to say that they offer a prima facie reason to suppose there is a problem with alien universals (and non-actual objects in general): the problem is that since we are in no contact with such entities, and they do not feature in our experience, we cannot talk about them or formulate truths about them. Whether this concern about the meaning of nouns and noun-phrases is a concern about "reference", and whether meaningful talk of non-actual objects involves "reference" to non-existent objects is a controversial matter, and is probably a matter to be settled in large part by a stipulation about how we wish to use the word "reference". Let me stipulate for purposes of this chapter then that I will be talking as if this is putatively a case of "reference" — though if your preferred usage of that term rules this out as a case of reference, feel free to take me to be talking about "reference*".

52 I will be talking about the relation of our words to the things named by those words, but I do not intend by this the view that nouns and noun phrases are all "proper names" (for my purposes "flying horse" and "green round hairy object" are phrases which name, even though they are not "proper names" in the way that "Aristotle" or "Pegasus" are. Neither do I mean to imply a "name" theory of meaning, in which all words name something (including words like "quickly" or "and"), or that all sentences are names (of propositions perhaps, or of truth-values). I am concerned here primarily with the meaning of nouns and noun phrases.
or some such. The issue is one about representation, rather than of reference *per se*. Similarly, for purposes of this chapter I will often speak as if there are non-existent objects, non-existent universals, and so on. Those that do not think such talk is legitimate are invited to paraphrase what I say by employing 'linguistic ascent', or contextual presupposition, or implicit operators of some sort, or whatever they favour: the points I make in this chapter do not rely on the appropriateness of neutral quantification or any form of Meinongianism. It is merely the idiom common in discussion of modality (and indeed in most philosophical discussions which employ possible worlds) of talking of *possibilia*, and sometimes quantifying over them.

An attractive solution to the problem of referring to non-existent or non-actual individual objects (or even actual objects not in any sort of direct causal or experiential contact with us) is that we can discuss those objects via descriptions. Provided we are in contact with general features of the world such as redness or hardness or rockiness (whatever the metaphysical status of these general features may turn out to be), it seems reasonable to suppose that we can talk of red hard rocks in other galaxies, or red hard rocks that could have come into existence last night but did not. We can also quite reasonably, it seems, talk about red hard rocks in other possible worlds (again, regardless of what metaphysical status such worlds turn out to have), despite our failure to have those rocks in an appropriately direct relation to us: we have the appropriate connection to the general features such objects have, and so can quite meaningfully use descriptions such as "the red hard rocky thing that...". In turn, our ability to use such descriptions meaningfully may allow us to use proper names to stand for objects that we are not directly in contact with (performing dubbings by using referential descriptions, for example, or by some other method).

The problem comes when we come to consider the general features themselves: the properties and relations. Our intuitions are that it is logically possible that there be universals other than those that there actually are (there could be another colour perhaps, or things might have had magic or spooky properties even if nothing does — it is logically possible that there be telepathic brain-waves or super-intelligent Martians even if in fact there are none, and so on). The problem is that perhaps we cannot individuate such non-actual but possible universals sufficiently using only universals that we are in appropriate contact with. Things could have been radically different from the way they are — there might have been no time, space, matter, causation and so on, but instead there could have been quite different properties and relations. But how are we to individuate these simply in terms of the general features we are in contact with? *Ex hypothesi* these alien universals are not universals we are in contact with ourselves, and they could well be so different that none of them can be distinguished simply on the basis of relations they have to non-alien universals — they may have very few relations to such universals, and more than one alien might have the same set of relations to our actual universals.
(more than one alien might be a type of charge, for example). Notice that this means that the problem of alien universals does not merely arise when we wish to talk about the alien properties or relations themselves: it also will come up whenever we wish to describe a merely possible object that possesses one or more of these alien universals. We seem to think that objects and creatures possessing properties or relations not found in the actual world are conceivable — we even conceive of them! An account is needed of how this can be done.

2. Denials That The Problem Arises

Before an attempt to deal with at least some of the strategies for allowing that there is talk of, and truths about, alien universals, let us first examine some of the positions which deny, in various ways, that there is a problem of alien universals at all: that we are unable to talk about or (by and large) produce truths about non-existent but possible universals.

The first strategy, often adopted by realists about universals, is to be Platonist about universals, in this sense: they take it to be the case that some universals that exist are un instantiated, and that every possible kind of universal exists. For example, there may be no ideal Republic to be found on earth (or indeed anywhere in Actuality), but nevertheless the property of being an ideal Republic does exist in the actual world. There are no missing shades of blue (at least in one sense) for such a Platonist: it may just be that some shades of blue are not instantiated by any entity. The Platonist can deny that there are or could be any alien universals: there are no non-existent but possibly-existent universals, because the range of universals that the more sparse of us take to be only possible the Platonist takes to be actual, but merely un instantiated. The Platonist can avoid claiming that there are alien universals, but nevertheless can agree with many of our modal intuitions about how things could be. Platonists can agree that there are worlds where very few of the actually instantiated universals are instantiated, and many universals completely beyond our experience are instantiated instead. It is just that those strange universals exist at this world just as much as the more everyday universals do.

Platonism is not really an adequate solution for the real problem that alien universals pose us, however, though the problem raises itself in a different form. The problem arises from a limitation on our linguistic resources: if we require experiential contact (of even the very weak sort which is offered to explain our naming of things only through descriptions which ultimately are sourced in our experience) as the basis of our ability to use names meaningfully, then causally inert un instantiated universals are in the same boat as merely possible ones: they cannot be described adequately in terms of the

---

53 Or at least every possible kind of pure universal exists. Impure ones, that is properties or relations which essentially involve particulars, (e.g. being to the left of Pegasus) might not exist even in Platonic heaven if the individuals in question do not.
(only instantiated) properties or relations that have entered into our experience. The problem of "alien universals" becomes, for the Platonist, the problem of uninstantiated universals. Platonists avoid even this problem, of course, if their view is combined with an epistemology that provides for some sort of experiential contact with uninstantiated universals: maybe through the direct awareness of transmigrating souls temporarily freed from their earthly fetters, or through a "Realm-O-Scope" inside people's heads that functions to provide us with direct awareness of uninstantiated universals, as some who believe in a faculty of Intuition seem to believe. For the majority of Platonists who have no such views, however, the problem is still pressing.

The same sort of problem raises itself for David Lewis's treatment of modality. The Lewisian does not have a problem with explaining how it is that there are truths about non-existent but merely possible universals and their distribution: because modal truths are cashed out as what is going on in a multitude of worlds of exactly the same sort as this one: saying that an alien is possibly instantiated but is not actually instantiated is much the same sort of claim as claiming that some object has a property that is not to be found on Earth: in both cases it is often just a matter of the object being at a different space-time location from us (though in the former case the space-time location is not connected to our space-time). The Lewisian does face the real problem however: in this case it is not non-existent nor non-instantiated universals that cause the difficulty, but it is non-actual(ised) universals that face the problem that the explanation of their actual naming in terms of things in our experience needs to be accounted for.

Of course, there is always the alternative of accepting that no alien universals are possible after all: or at the very least, no statement about the possibility of such alien universals is true (as this problem seems to be one to do with our language, and not necessarily with modality as such). Some theorists about modality have thought this (Armstrong 1989 is an example), but it seems a counsel of despair unless no other option is available to us. For one thing, we take ourselves to be able to talk about alien universals on a regular basis (fiction is full of them), and even in the present discussion I have repeatedly purported to talk about such entities. Furthermore, if it is admitted that there are possible worlds which contain fewer universals than this one, (and so that relative to those other possible worlds, the actual world contains alien universals), then it must be admitted that some worlds are such that some universals are alien relative to them, and so it must be admitted that it is perfectly possible for a world to be such that some universals are alien relative to it. Since this sort of thing is possible in general, it would be strange if we could rule it out immediately in the case of the actual world. Of course this last consideration is only offered to motivate an examination of alternatives which would allow us to truly speak of alien universals — it is not a good enough

54 This point is made in Lewis 1986c on p 159.
argument to show that we must be able to speak of alien universals. After all, it might be thought that the actual world may have a special status compared to the others, and part of this special status may be that it contains the most universals that there could possibly be. It might have this feature in virtue of being the only world that exists, or through its being the only ground of truths, or some other special distinguishing feature (see Armstrong 1989, p 56-57). The "no possible aliens" view is not an untenable one, but a rival view that accords more with our modal intuitions might be preferable if it is otherwise plausible.

3. Explaining the Caveat “not constructible out of existing universals”

The Platonist and the Lewisian, while they do not face the problem of alien universals, do both face analogous problems: the problems of ensuring that we can talk of things which seemingly we cannot either name directly or refer to via descriptions cast in terms of the properties and relations of objects that we are acquainted with. The Platonist and Lewisian should both find paraphrases of the following discussion useful then: the Platonist just has to read “uninstantiated universals which are not constructs from instantiated ones” where I talk of alien universals, and the Lewisian just needs to read “unactualised universals which are not constructs from actualised ones”. With this in mind, let me continue.

In this chapter so far, I have defined alien universals not merely as those universals which do not exist (or are not instantiated, or are not actually instantiated) but also as those universals which are not constructible out of the universals which do exist (or which are instantiated, or which are actually instantiated). The time has come to explain what this caveat amounts to. The caveat is not of my devising — it is part of the standard definition of this sense of “alien” as it is found in, for instance, D.M. Armstrong’s work (see e.g. Armstrong 1989 p 54). The addition of this caveat allows those who reject alien universals to partially satisfy our intuitions that there could be properties or relations other than actual ones, but as we shall see it does not deliver all that one might wish. The reasoning behind the caveat, however, has features which can be used to construct a more satisfying answer to the question of how we might talk of the truly alien universals. After producing the more satisfying answer, I will show how the more satisfactory solution deals with the objections based on the problem of alien universals that Lewis has raised against Linguistic Ersatzism and the objection raised by Armstrong against alien universals.

The most straightforward class of properties and relations which are not instantiated but might have been are certain conjunctive properties and relations. When something is red and is also round it therefore has the property of being red-and-round (assuming that redness and roundness are genuine properties). Some existing properties and relations are not found together — but there is little problem in describing situations
in which they are found together using descriptions built up from components which we are appropriately connected to. Many things have had the property of being diamond, and at least one person has had the property of having the shape of Marilyn Munroe — but it is plausible to think that nothing in the actual world which is diamond has ever or will ever have had that shape (though whether future high-tech sculptors will create such a thing is probably currently unknowable). Suppose that no actual thing is both diamond and has the shape which Marilyn Munroe had. Nevertheless, it is perfectly straightforward to describe something such that it has both, and so describe something which has the conjunctive property of being-diamond-and-having-the-MM-shape. Descriptions of non-existent conjunctive properties and relations can easily be constructed, and so these non-existent properties and relations are the first category to not face the problem which seems to face the alien universals. Conjunctive universals whose conjuncts exist are the first category of non-existent universals which do not share the problem of aliens. Slightly more contentiously, there is another category as well.

Armstrong admits that there are some other merely possible but non-existent properties and relations, but claims that all of these are (or would be) structural properties and relations that have only actual universals as their constituents. And in arguing for this he points out what seems to be a perfectly acceptable method of providing uniquely identifying descriptions for some non-existing universals. Some people (especially Lewis 1986a) have objected to structural universals, and it may be thought that Armstrong’s partial solution only works for those who are Armstrongian realists about properties and relations. After I have outlined Armstrong’s strategy, however, I will argue that it has general application and does not rely on some of the controversial specifics of his account of structural universals.

Armstrong claims that some universals are structural — that is, they are properties (or relations) of structured states of affairs consisting of instantiations of various more simple properties or relations. One example is the property of Being Methane. To instantiate this property is just a matter of being a structure consisting of four hydrogen atoms covalently bonded to one carbon atom. The property of Being Methane, for example, is nothing over and above a pattern of instantiation of the properties of Being

55 Those who have quibbles about the example — worrying, for example, that the properties I ascribe are not genuine properties, or who are pretty sure there will be MM-shaped diamonds somewhere in the actual world are, of course, invited to substitute a more congenial example. And of course Munroe possessed different shapes at different times, so only one of the representative shapes should be chosen.

56 Also included among the non-alien will be conjunctive universals whose conjuncts include such structural universals. Indeed, any conjunctive universal whose constituents are non-alien will itself be non-alien.

57 I will employ capitals when I am employing the name of a property — so, for example, hydrogen atoms have the property of Being Hydrogen, and electrons have the property of Negative Charge.
Hydrogen, Being Carbon, and mereological relations and the relation of Covalent Bonding. (In turn, it seems obvious that the Hydrogen and Carbon properties and the Covalent Bonding relation are themselves structural — they can be reduced into patterns of various sub-atomic states of affairs). Structural universals might even be the only sort of universal that there are (Armstrong 1989 p 113). Structural universals are instantiated when and only when the structures they supervene on exist — it would be nonsensical to think that the Methane property (defined in terms of the existence of the structure involving hydrogen and methane which we know methane to have) could fail to be instantiated when there was a case of just four hydrogen atoms appropriately bonded to just one carbon atom, or that the Methane property could appear without that pattern of instantiation of the simpler universals. They can be exhaustively described in terms of the simpler constituents they have, and those constituents’ arrangements (though the structural universal may stand in causal relationships or other external relations that its constituents do not).

It seems that in the case of such structural universals, we can produce an individuating description of them in terms of the universals which are their constituents and the arrangements of those constituents. I can define Methane in terms of other universals, and it seems I could do so even if there were no methane — provided I was appropriately acquainted with hydrogen, carbon, covalent bonding, and so on. Similarly, let us suppose that there is no actual golden mountain. Nevertheless, I can define the property of Being a Golden Mountain in terms of size properties (being of at least such-and-such height, being at least so-and-so width, having a shape of blah, etc.) and the property of being gold. “Being a Golden Mountain” is then a non-existing property — yet we have a method of picking it out. This is the strategy which is suggested by Armstrong’s brief remark to the effect that the “missing shade of blue” would be a structural universal composed of actual constituents and thus not something that need be ruled impossible by his theory (Armstrong 1989, p 56), and by his account of structural universals as being supervenient on states of affairs appropriately instantiating their constituents (p 113). This kind of strategy might also have been part of what David Lewis had in mind when he defined alien properties and relations as not only failing to be actually instantiated but also not being “analysable as a conjunction of, or as a structural property constructed out of, natural properties all of which are instantiated by inhabitants of that world” (Lewis, 1983, p 364)\textsuperscript{58}.

\textsuperscript{58} Strictly speaking, Lewis was defining the relation of “alien to a world” rather than the notion of a non-relative distinction of the alien vs. the non-alien, but this is not important for present purposes, as Armstrong means by “alien” what Lewis would mean by “alien to the actual world”. Note that this means that Lewis’s use of the term “alien” would not be my use. He introduces the notion of alieness in the context of a discussion of materialism, and he obviously would want to include in the obviously material worlds which differed from a given material world only in virtue of their having instantiated in them a property or relation not instantiated in the given material world but which was merely constructed
Some might be tempted to reject the strategy of saving meaningful talk about universals which do not exist but are constructible from existing universals. One ground for this rejection would be the belief that structural constitution either makes no sense or is not metaphysically required (see Lewis, 1986a, for a classic example of an opponent of Armstrongian structural universals). There is no reason why the relevant class of properties and relations have to be considered as structural for something like the solution outlined to work. Take again the property of Being Methane. One does not have to accept Armstrong’s specific account to think that Being Methane (or, for simplicity here, Being a Molecule of Methane) is nothing more than something’s being made up of (just one) carbon atom and (just) four hydrogen atoms to which the carbon is covalently bonded. Furthermore, one does not have to buy too much of Armstrong’s metaphysics to believe that this is the only way that something could become methane (or at least that it is the only way that something could become the kind of methane which actual molecules of methane fall under). Provided one thinks these things, and thus thinks that a description in terms of carbon, hydrogen etc. uniquely picks out the property of being methane (or one of the determinates of that property), then one should agree that that description could meaningfully pick out the property of being methane (or, again, one of the properties of being methane) in worlds where there were carbon atoms, hydrogen atoms etc. but no methane. Similarly, the style of solution works to allow us to talk about the property of Being A Golden Mountain in virtue of our being able to discuss the property of being golden and various size properties. It does not matter for current purposes what specific account is given of the connection between the “simpler” properties and the “more complex” properties is, or at least would be.

An example of a theory that would provide the resources to allow for this kind of “construction” without taking it to be a matter of structural universals would be David Lewis’s own. On his account, the property of being methane, the property of being hydrogen, the property of being carbon and the relation of covalent bonding would all be sets (see Lewis, 1986a). The mereological “relation” of part-to-whole is slightly trickier — it is not to be cashed out in terms of being a set, but is rather to be taken as a primitive dyadic predicate. Still, we certainly have epistemic and descriptive access to mereological connection, even if its metaphysical status is a little unusual. Since we can define the membership conditions for the Methane Atom property (or at least a Methane Atom property) in terms of the Hydrogen Atom property, the Carbon Atom property, the Covalent Bonding relation and the mereological connection, we can provide an out of properties or relations to be found in the given world.

59 The part-whole “relation” cannot be taken to be a standard set-theoretic relation for Lewis on pain of class-theoretic paradoxes given Lewis’s view of the connection between mereology and set-membership, and on pain of vicious circularity given the account offered of reducing class theory to his “framework”, which includes mereology. See Lewis 1991 and 1993 for details of how mereology is connected to set-
individuating description for the Methane Atom property even when we have no access to the Methane Property (or any of its members) directly. A matter of set-theoretic definition and perhaps invocation of mereology should be enough for us to individuate Being a Golden Mountain in a similar way.

Of course, this strategy for allowing for talk about merely possible universals that are constructed out of simpler actual universals will not be a completely satisfactory answer to the problem we are faced with, as there will still be the problem of aliens: those universals that are not constructions of simpler universals. We would like to say that there could have been other simple properties besides the ones that there are, and we would like to say that there are possible properties that are constructed at least in part out of properties and relations that are not themselves constructed out of actual constituents. The method of dealing with those universals which do not exist but which can be constructed from those that do exist does, however, provide us with a clue for how to accommodate the alien universals.

4. How to Represent Alien Universals

To try to formulate descriptions of the alien universals (those properties and relations that are not instantiated, and cannot be “constituted” out of actually instantiated properties and relations in any way), our instinct is to employ the sort of predicates which are often used to form descriptions of actually instantiated properties and relations. We might say of a putatively alien universal that it stands in such-and-such nomic relations, that it is instantiated by so-and-so things, and that it is distinct from (i.e. not identical to) all of the actually instantiated universals and all of the universals that can be constructed from them, and so on. We can attempt to state the -adicity of the alien universal in the obvious way, perhaps state which genera it falls under, and so on. No-one disputes that we can meaningfully and truthfully employ these sorts of predications in general, so why is it that they cannot possibly apply to the alien? If they could, of course, then our problem would be solved, since we could talk about aliens through providing descriptions of them, just as we can do for more run-of-the-mill non-existents.

An interesting question is the metaphysical status of such predications. Do they ascribe properties and relations to universals (or between universals and other things), or is there some other way in which these predications become true? Elementarists (e.g.

theory for Lewis.

60 Well, of course some philosophers do dispute just this, but I am assuming for the purposes of this chapter that the problem of alien universals is a particular problem for aliens. Those philosophers who think it is senseless to talk of alien universals and predicate various things of them because it is in generally senseless to talk of universals and form descriptions of them will not be having their position considered here. Not only must explanations come to an end somewhere, they must also start somewhere.
Bergmann 1957), for example, will want to claim that such predicates, if they are truly
applied at all, are not applied in virtue of further properties or relations. Armstrong will
apparently wish to deny that truths about instantiation of universals and non-identity are
to be explained by the existence of relational universals between universals and their
instances in the one case (Armstrong 1978a p 108) and between universals in the other
(non-identity may fail to be a relation on two counts for Armstrong — the first is that it is
seen as a negative predicate, and the second is that identity is an internal relation on
Armstrong’s account (see Armstrong 1978b p 86-87), though at least the nomic
connections are meant to be in virtue of higher-order universals (Armstrong 1978b p
149). My view is that these predicates are all associated with (higher-order) universals,
but I do not think that the opposing views need have difficulties in applying the
predications to the aliens. Regardless of whether the predications are true in virtue of
relations, these theorists must think that they somehow get to be true sometimes and that
we have some kind of epistemic access to truths involving such predications. Provided
that we have understanding of the predications, regardless of how we come by such
understanding, we should be able to understand some descriptions which are partly cast
in terms of such dyadic predicates. So the metaphysical status of such predications need
not worry us overmuch.

One thing that is a problem is providing a unique description of an alien universal
(that is, a description that applies, or would apply to only one alien universal). If things
like the nomic connections between universals and whether or not given particulars
instantiate them are quite contingent matters, then it is reasonable to suppose that if one
alien property stands in a given network of nomic relations and is instantiated by a given
collection of particulars in one world, then there is another possible world where another
alien property stands in those same relations to those same things. And while non­
identity from the actual (and constructible from the actual) universals will presumably be
something which an alien universal will have necessarily (assuming the facts about the
identity of universals are necessary, which is a common enough assumption), this will
not help in uniquely picking out a possible satisfier of a description, as all the alien
universals are non-identical to the actual and constructible-from-actual universals. Is this
lack of a uniquely identifying description fatal to the task of talking about or naming alien
universals? I think it is for some processes of naming, which do require a uniquely-
designating device (such as a definite description) to support the name. “Dummy names”
can still be supported in the usual way, it would seem: just as “Take an electron. Call it
A...” is a meaningful way to introduce the “dummy name” ‘A’ without providing a
unique description that distinguishes A from any other electron. Without real names for
alien universals, it seems that we are unable to make particular claims about one alien
rather than another. But surely we are still able to make general statements of the form
“There is a universal, distinct from all the actual and constructible-from-actual universals,
that has these relations to X, and those to Y, etc. etc.”. What reason do we have to suppose that these kinds of claims must be meaningless or necessarily false?

One bad reason would be that without the capacity for singular reference (or “picking out”, or whatever the equivalent of reference is when we are dealing with things that do not exist) we do not have sufficient resources to number the aliens, or say that they are distinct from one another. This might be thought to be so because claims about the number of aliens (if we try to say, as surely we want to, that the number of aliens is greater than one) involve claiming that they are distinct from one another, which might give people worries about circularity. This is because to make it true that an alien is distinct from all of the others, we must mention the others — but, it might be thought, we have no authority to suppose that the others are represented to have determinate identities until we can at the very least say that they are distinct from the first. Fortunately, we know how to meaningfully introduce talk of objects which are inter-definable without risking circularity: it is the familiar technique of using Ramsey sentences with more than one existential quantification. The worry that some of the descriptions we wish to say possibly apply to alien universals involve reference to other alien universals need not worry us, then, as we know inter-definability is not something to worry about in this sort of case, and in particular it does not lead to any vicious circularity.

There is another worry about alien universals that is based on the observation that the statements about them can only be of a general form. This worry, while more plausible than the definitional circularity worry, can also be dealt with in a fairly straightforward way. This is the worry suggested by a comment of Armstrong’s (Armstrong, 1989, p 55-56). Armstrong’s worry is summed up in the following quote:

“What is possible could be. Something merely determinable could not be. To be is to be determinate.”

The worry sounds like it is something like this: it seems that at best we can only express the possibility of general facts about aliens. But the general facts could only be possible if specific facts about aliens were possible. So the general facts cannot be possible either. Put like this, it is clear that the premises are true. (The general can only be made true by the particular, after all). But equally clearly they do not support the conclusion. Just because we cannot express any of the particular propositions does not

61 Skyrms mentions the use of Ramsey sentences to do something along these lines in his (Skyrms, 1981). While it is obvious to those familiar with Ramsey and Carnap sentences that they can be used to avoid some of the fears of circularity which come with inter-definition, an example in another context of how they can be used to avoid circularity can be found in (Lewis, 1994).

62 In fact Armstrong is not primarily worried about the representation problem, but the specific problem with alien universals which arises for his theory qua combinatorial theory. That is a legitimate worry for his theory, and the remark I quoted need not and perhaps should not be seen as making the mistake I claim one would make if one had the worry I use this quote to illustrate.
mean they are not possible. In fact, surely we would want to say that in any possible world where there are general facts true of alien universals, there are particular facts true in virtue of which the general facts hold. The fact that we cannot represent something cannot make it impossible: after all, the conjunction of all the truths cannot be uttered by we merely finite creatures, but it is not plausible that the conjunction of all truths is an impossible proposition.

5. The Problem for Linguistic Ersatzism and Modal Fictionalism

Can Armstrong’s worry be modified so as to be more worrisome? It would be more worrying if we could argue that we must be able to state or represent a proposition, or at least a certain class of propositions, in order for them to be possible. Then the fact that we could not state or otherwise represent the particular facts might make plausible that the general facts could not therefore be possible after all. This is the plight that the linguistic ersatzers and representational abstractionists in general apparently faces, and this worry may explain Lewis’s charge that linguistic ersatzism cannot deal adequately with the “outer realm”. Lewis’s arguments will be examined in more detail below, but first let us examine whether the linguistic ersatzer needs to take Armstrong’s worry as damning the project of allowing for the alien universals.

Linguistic ersatzers (see Lewis 1986c, 142-165) take possible worlds to be maximally consistent sets of sentences, or something at least broadly analogous — worlds are taken to be made up of things which are representational and have their representational properties in part through syntactic structure. In Lewis’s words, the worlds should be made up of “a system of structures that can be parsed and interpreted” (Lewis, 1986c, p 144). According to a linguistic ersatzer, P is possible iff a possible world represents that P. So for a linguistic ersatzer, if s/he accepts that for general claims to be true there have to be some corresponding particular claims true as well, a possible world would have to represent some specific claims about aliens for the kind of general claims we want to be represented as true by that world. For example, if we want a possible world to represent that some alien universal is instantiated, it must also represent, for some particular alien universal or other, that it is that alien universal which is instantiated. So the linguistic ersatzer needs to be able to claim that some actual representational structure can represent a specific alien universal. Linguistic ersatzers also are characterised by taking it that the syntactic structures which count as worlds are to receive their interpretations in the standard way that interpretations are assigned, so that

63 Those that think that propositions must be entertained in order for them to be able to be true might take the force of this argument to flow the other way — there must be an infinite, omniscient being in order for the conjunction of all truths to be itself true — an Idealist proof for the existence of God! This takes us rather off the topic, I fear.
they are bound by the same broad constraints on meaningfulness and representation as
e.g. sentences of a natural language. In other words, they do not rely on “magic”
representation or meaning. Relying on “magic” representation or meaning solves the
problem of representing aliens, for it just take it to be a primitive fact of the matter that
they do so appropriately. Ersatzism only faces a problem if we take the representation
which the ersatzer’s structures do to have broadly similar features to the representation
which we mere mortals manage.64 Thus not all representational abstractionists are
affected — those who postulate “magic” representation are typically not affected by this
problem. Since the problem is specifically for those representational abstractionists who
fall under Lewis’s “linguistic ersatzer” classification, I will refer to them as linguistic
ersatzers for the remainder of this chapter, despite the reservations I expressed about the
term “ersatz” in chapter one.

The modal fictionalist is in the same boat. Modal fictionalists claim that there are
no non-actual possible worlds, but we can employ a fiction to the effect that there are
such worlds when we engage in possible-worlds talk. For a modal fictionalist, a claim
about the possibility of alien universals will be true if and only if the modal fiction
represents that such an alien universal is instantiated at some world.65 So if the modal
fiction does not represent particular claims about alien universals, such claims are not
possibly true, by the lights of the modal fictionalist. For simplicity, the following
discussion will be cast in terms of linguistic ersatzism, but the same points apply equally
for the modal fictionalist (unless such a fictionalist avails herself of magically representing
fictions, as the magical ersatzer avails himself of magically representing worlds), since
she too will require an actual representation to represent every possibility.

The linguistic ersatzer seems to be in trouble. But again, the trouble is more
apparent than real. For consider the linguistic ersatzer with her representation: call it the
“world-book”. The huge Ramsey sentence stating what the alien universals66 are like

64 As Lewis points out (p 161) a Platonist of the sort I described in section 2 has resources which will
help here — for if all of the alien universals exist, the Platonist can specify that each is its own name,
and thereby have a stock of representations which represent each alien. For the sake of this chapter, I will
go along with Lewis as taking this Platonism to be another method of employing “magic” representation,
and so not to be a method open to the linguistic ersatzer, or the fictionalist constrained by the scruples of
the linguistic ersatzer in constructing her fiction.

65 The modal fictionalism of Armstrong 1989 is slightly more complicated — Armstrong’s modal
fiction represents that there are as many world-fictions as worlds, and represents of each of those world-
fictions that they exhaustively represent a world (Armstrong does not identify these fictional world-
fictions with possible worlds, since he takes worlds to be concrete — see p 50 of Armstrong 1989). This
added complexity does not alter the fact that his view would face this challenge if he admitted that alien
universals were possible, since his modal fiction must, for each possibly true claim about aliens,
represent of that claim that it is contained in one of the (merely fictional) world-fictions.

66 “world” is ambiguous, as I have discussed in the first chapter — it can refer either to a maximal state
of affairs, or cosmos, or whatever (the sense in which abstractionists believe that there is only one world),
inter-defines the alien universals, and inter-defines the alien universals and their instances and the universes which contain those instances. Such a massive Ramsey sentence will entail that some universes contain instances of alien universals (which is, of course, a general claim). And, for each universe which a claim of the form “an alien universal is instantiated at...” is made true, the Ramsey sentence will entail that those universes have in them instances of alien universals. In the English translation of this huge Ramsey sentence, something like “Property $X_1$ is instantiated by individual $y_1$ at universe $w_1$” will appear in it somewhere. But $X_1$, $y_1$, and $w_1$ are variables which will not be able to be replaced by any more than dummy names in such a case, because there will be other such triples of things with the same relations to any actual things, but will differ only in their relations to other alien universals, their instantiation in other universes, and so on. This inter-definition will not only be an inter-definition of alien universals, but one of certain universes and individuals as well. For two universes which differ only in virtue of it being the case that wherever some alien universal $A$ is instantiated in the first a different alien universal $B$ is instead instantiated in the second will be different in a respect that can only be specified through claims about alien universals.\footnote{Unless the Identity of (both qualitative and relational) Indiscernables is false for possible worlds, in which case there are two ways the worlds can be different — their difference in respect of alien universals and their “bare distinctness”, however that is to be cashed out. I suspect their “bare distinctnesses” would have to be an additional reason why they would have to be Ramsey inter-defined, were there such bare} The same may well go for objects in a universe which differ only in that the first is an instance of one alien property and the other is an instance of another of the alien properties. If it is true that there is no necessary upper bound on the number of alien universals which are possible, the inter-definition will be \textit{very} big — for there may well be at least $2^n$ universes where $n$ = the number of alien universals, and there may well be even more \textit{possibilia} than that. One thing the ersatzer has to accept is that the representation of the way things could be must be \textit{very} large. This whole infinite inter-defined mess will lack an “entry point” — a way of specifying solely with respect to the things we can make particular reference to some distinguishing feature that one of the things in the mess has that no other possesses. Yet we do not appear to have a situation where the description of what is going on entails that a general claim is true at a world without a required specific claim being true. What could be going on?

The answer to this will first require a distinction between particular claims and specific claims (the selection of terminology being somewhat arbitrary). Particular truths rely on particular reference (or whatever you want to call “picking-out”, in this context).

or it can refer to the things which play the theoretical role of “possible worlds", as it is normally conceived of. Ersatzers believe that there are many of the second, and that they are abstracta of various sorts, whereas fictionalists do not believe in them, but believe that according to the Modal Fiction they exist. To prevent confusion, in this chapter I will use the word “universe” to capture the first sense, and “world” for the second.
They rely on proper names, or predicates that have become meaningful through some sort of not-merely-descriptive mechanism (e.g. predicates that are made meaningful through some sort of causal chain from the relevant property or relation, or by means of a referential description of a property or relation, or of the predicates' actual extension, or whatever). The particular claims I have been contrasting with general claims (those claims that are not, at least in a relevant respect, particular). Specific claims, on the other hand, are about an individual individuated in a sufficient manner by some means. Particular claims would then often be specific ones as well, but there will be exceptions: the expression “Something to the left of Plato” is particular to the extent that “to the left of Plato” is a predicate which is meaningful in part through the causal chain associated with the name “Plato”. It would not be specific, however. On the other hand, specific claims could be made entirely through description mechanisms (through claiming something about a thing for which a unique description is provided). It is the specific claims which are more properly considered the opposite of general claims. The class of general claims and the class of particular claims can even overlap: consider, for example, the claim “Some art produced after Picasso is decadent” — it is general, and cannot be true unless a more specific claim is true (it must be true for some given piece of art that it is produced after Picasso and is decadent), but it is also particular, since it relies on a “picking out” involving particular reference to Picasso. Now, specific propositions must be true whenever general ones are, but there is no reason in general to think that there need be particular reference techniques available to express such specific propositions as particular ones, or alternatively that propositions represented with particular representation mechanisms must be true whenever general propositions are. The worry about merely general representation need not worry even the linguistic ersatz:er: for the world-book represents that for any given universe, there is one and only one alien universal in each universal-role in that universe, and furthermore through the huge inter-definition the world-book states that the alien universal in question is distinct from all the others, has such-and-such roles and attributes in other universes, and so on. The world-book represents that there are specific truths for each world, even though it is true that all the representations made by the world-book, or that are contained in the world book, are all non-particular. The world book consists of a specific claim about the transfinite totality of the collection of universes, individuals in those universes (including, presumably, individuals not found in the actual universe) and universals, both the alien universals and the more ordinary variety.68

68 This is assuming that for simplicity we take it to be the case that the world-book is making only one claim (this makes using Ramsey/Carnap devices more straightforward). The use of dummy-names, anaphora etc. will allow the world-book to be making many claims about the worlds and their contents, which might be a more intuitive way of looking at it. This complication is not a difficulty in principle
To set up an example which might help to make clear what would be a legitimate worry about specificity vs. generality and what would not, let us consider two (very cut down) representations. The first represents that a certain number of objects exist, and also represents various general claims about the collection of objects (they are such that one is red, one is round, three have volumes of greater than a cubic metre, and so on). This could not be a complete representation of a way things could be. This is so because if all of these claims were true, there would be further things that would be true and which would not be entailed by the claims in the representation: namely, the truths about which thing is the red one, which ones are larger than a cubic metre, and so on. Presumably, the representation could be made true by different possible states of affairs: some in which the red thing is also the round thing and others where the red thing is not round. Consider now a second representation, which represents the same things about the number of objects and the general truths as the first representation, but in addition tells us which of the objects has which property, but only through using more general statements. So it might tell us that one of the objects was red and square, and had certain dimensions, spatio-temporal location, and was composed of a certain material, and so on... including the assertion that they were all the features (or all the “basic” features, or that the intrinsic properties mentioned were all of the intrinsic properties possessed by the object). It could then go on to do this as many times as it says there are objects, and so give a complete qualitative description of the objects and their relations. This second representation, I claim, could be detailed enough to completely represent a way things could be (or, given the linguistic ersatzer’s account, there are representations of this second sort which are ways things could be). In asserting this, I reject a form of haecceitism. This form of haecceitism rejects both of the following claims:

1) A thing’s identity (being, existence, substance or whatever) is given by, or supervenes on, facts about which universals it is an instance of, or is part of an instance of, or is ordered in an instance of.

2) There are “individual properties”, not supervening on any more general universals or instantiation of universals, and it is only these individual properties which give a thing’s identity (being, essence, substance, or whatever).69

Hopefully, though, such a haecceitist will at least accept the following even if they reject (1):

69 (2) might be thought to be implied by (1), and so not worth separate mention, but it may well be controversial whether “individual properties” are universals (or genuine properties or relations) at all, so I mention it separately. It is also worthy of separate mention for it is recognizably a doctrine which is sometimes associated with the name “heccaetism”.

if one wishes to include it in one’s linguistic ersatzer account, so I will not bother to deal specifically with this variation of ersatz world-books in my discussion here.
3) A thing’s being, essence, substance, or whatever, can be given by its modal qualitative identity conditions — what its complete qualitative description (including complete specification of what relations it is ordered in an instance of) is in each possible universe.

I happen to think (1) is true of mere particulars but perhaps false of universals (and my example relies on (1) being true of particulars), and that something like (3) is more likely to be true of universals, and (2) is, though attractive to many, an ontological mistake. However, if a haecceitist takes (3) to be true of both particulars and universals, s/he could reject my simpler example and yet allow that the linguistic ersatzer has represented enough so that the world-book does not suffer the problems of my first example representation, as the world-book has what my second example representation does not: a specification of modal qualitative identity conditions (albeit through Ramsey inter-definition, as the qualities as well as the objects are being represented through their modal identity conditions). So the worry about particular representation, suggested by the quote from Armstrong, as applied to linguistic ersatzism either rests on a confusion or is simply just inapplicable. (Unless of course (3) is rejected. But why should anyone, let alone a linguistic ersatzer, reject (3)?) The Armstrong worry is not worrisome after all, because showing that particular “reference” to alien universals is not possible is not to show that singular “reference” to them is impossible. Not only was the original argument which I used to set up the worry not workable as it stood, the implicit premise of the impossibility of there being specific descriptions of alien universals was, it seems, not even necessarily true for the linguistic ersatzer.

Some may still feel a residual worry. Particular claims can be made about those things which are not alien, but such claims cannot be made about alien universals, or things which instantiate alien universals. However, particular claims in the mouths of inhabitants of universes containing alien universals would be true if that universe were the actual one. For example, if there were a sixth sense that picked up on the pattern of instantiation of an alien property through some direct causal connection, then if people decided to call the objects in which that property was instantiated “Ugs”, then through a referential description, or some causal-chain involving dubbing procedure, such people could make the word “Ugness” name the property in question in the same general kind of way in which we name actual properties. “Ugness is instantiated” or “Ugness causes sensations of Ug” would then be particular claims which, were they made by these inhabitants, would be true. Yet we cannot use “Ugness” or “Ug” in a particular manner, as we have no way of particularly identifying Ugness (or the class of universes in which Ug is instantiated, for that matter). Would there then be propositions that are actually
inexpressible but would be expressible? Does this show us that after all that the ersatzer’s world book cannot, even with Ramsey inter-definability, represent ‘the universes of Ug’?

The linguistic ersatzer can plausibly claim that it does not: it is just that the Ug-worlders would have a different way of expressing the claims which the world-book makes about their world. “I am here” in my mouth expresses the proposition that “Daniel Nolan is in his office in the Coombs Building”, but the first way of expressing that proposition is one that only I could have used successfully to express the same proposition as the second sentence expresses. Ug-worlders can uniquely pick out their universe with the phrase “the universe I am in”, or “the universe that exists” (understood in some sense), whereas such indexical phrases in our mouths pick out the actual universe. Given that plausibly it is indexicality (or an analogue) is what is providing them with the power to make particular claims about their universe, or themselves, or the things which impact on themselves, we should think that actual representations can in some sense be representations of all the things which they can represent — it is just that particular representation is a different kind of way of representing that they can do of themselves and their surroundings but that we cannot (but see p 117 and following for further discussion of this issue). The situation is symmetrical of course — if there is some property in the actual universe which is not found in the Ug-worlders’ universe and cannot be structurally built up from universals found in the Ug-worlders’ universe, then we will have the power to particularly represent things about our universe that representations in the Ug-universe will not.

Lewis, in considering a strategy for producing multiplicity of representations simply by pairing representations to integers, says this pairing strategy is simply introducing an “irrelevant multiplicity” of representations (Lewis, 1986c, 163). After all, all we would have through this strategy is many ambiguous, multiply-realisable representations of universes instead of only some such ambiguous representations. So I agree that irrelevant multiplications of representations would not help the linguistic ersatzer’s cause. So why is my multiplication of representations not irrelevant? To answer this worry specifically: the world-book not only contains many representations that differ not at all in intra-universe matters (that is, do not differ _qua_ a general description only of the pattern of arrangement of objects and universals in that universe): it contains representations that the things represented by each of these representations are different (this through something analogous to Ramsey sentences containing claims that various things are not identical to other things, or that there are universes where both exist but have different properties and relations, or whatever). So the multiplicity of representations are different, in the context of the world-book, because the world-book itself represents that the objects they represent are different. To put it bluntly — the relevance of the multiplicity of representations (that are in one sense the same sort of representation) is that the world-book says the things which each represents are different.
from the things that all the others represent. A representation that two variables represent different things is enough to make it the case that two representations which vary in that one employs the first variable and the other employs the second are in fact different representations of how things are, provided that the latter two representations have their variables in the scope of the bindings of variables in the former representation. So this multiplicity escapes the irrelevance charge.

6. Making Sure the World-book Represents the Actual Universe

I have argued that one worry about the world-book's representative capacity is unfounded, but there is a similar worry which does need to be addressed. The world-book represents an infinite number of possibilities, but there is not yet any explicit guarantee that it represents the actual world as one of them. Some accounts of the recursive specification of the world-book, and the limits upon what worlds it represents as possible, will have ensured that the actual universe is represented, but in the absence of some such guarantee dropping out of the specification procedure the guarantee that it represents the actual universe must be added explicitly — for an ersatz or fictionalist theory of possible worlds cannot hope to be adequate unless what is actual is possible as well. There are a variety of methods one might wish to employ to explicitly ensure this, and I will not try to exhaustively account for such methods or even detail any with the rigour which would be required to properly set out an account of the world-book — that is the task for one constructing or defending a linguistic ersatz theory as a whole, rather than for this chapter, which is merely dealing with the question of alien universals. However, I will briefly discuss the sort of thing which might be done, if only to persuade the reader that addressing this challenge is orthogonal to the main challenges addressed by this chapter.

I will assume that the linguistic ersatzer or fictionalist has the theoretical resources to allow for a structure (of the same general kind as the world-book I have been considering) to represent the actual world in all of its detail. Explaining how it might be that such a (presumably abstract) structure can do this is no easy matter, but if the ersatzer or the fictionalist cannot even adequately explain how abstracta can represent the actual world, then their projects cannot get started in the first place. The first step in adding the representative content necessary for the world-book to represent the actual world is to add a clause to the world book to the effect that one of the universes represented is the actual universe. Unfortunately, this is not yet enough to give us what we want. For the mere addition of the analogue of "and one of the universes is the actual one" does not ensure that the actual universe is represented by the world-book to be as it actually is — it is perfectly consistent with all that is said by the world-book so far that the actual universe is one populated by unicorns, blue swans, vital spirits, and who
knows what else. We also need the world-book to represent the actual world as it in
fact is. Fortunately, this does not have to be much of a challenge: for we have granted
for the sake of discussion that an object like the world-book can represent the actual
world faithfully, so we need only employ that representation technique to add that the
actual universe is that way.

Is it enough to simply add these clauses? It might not be if the ersatzer insists on
explicit representation — for while what has been specified entails that one of the
universes is the actual universe (and is the way the actual universe is), and this plus facts
about what is consistent with the rest of the world-book implies that some (or all) actual
objects are found in other possible worlds, and that actual universals are distinct from all
alien universals, and so on, it might not say this explicitly, and the ersatzer wishing for
her world-book to say everything it says explicitly (perhaps so that they do not need to
rely on modal notions such as implication or entailment) may need to do more. This
should not be difficult in principle — it simply requires the world-book to say more about
the connection between the actual universe as represented and all of the other universes
represented.

When specifying how the clauses about the actual world must be added, a
decision must be made about the modal status of the world-book at this point — will the
world-books at other worlds represent those universes as the actual universe, or will they
represent this universe as the actual universe? I think they must represent those universes
to be the actual universe, for otherwise, for example, some worlds will need to be able to
represent universals which are alien from their perspective (namely, those universals
which exist at this world but which do not at that world, and cannot be constructed from
the universals which are found there), and furthermore represent them directly, in the
manner which our world-book does. If this could be done then my appeal to the inter-
defined world-book is a waste of time — but if, as I suspect, it cannot, then the linguistic
ersatzer is stuck with the problem of representing alien universals all over again, but
without my help available. Fortunately, I think the conclusion that other-worldly world-
books represent universes other than this one to be the actual universe is the correct one
on independent grounds — for the context supplied by each of those universes is a
context in which it is correct to assert that that universe is the actual universe.70 This does
not mean that the theory says that which world is the actual varies from world to world —
for the correct world-book is our word-book, and it says that this universe is the actual
universe once and for all. How other-worldly world-books would represent is neither
here nor there (or in other words, what our world-book would represent in other worlds
is not the question of how it does represent the universes other than our own).

---

70 See for example the treatment of “actually” in Kaplan (1978). The same indexical or demonstrative
treatment would apply to the phrase “the actual universe”.
The practical importance of having the representation represent that the actual universe is one of the universes, and is as it actually is, is great. For without this, almost any modal claim we care to make which involves particular representation is in trouble. Take, for example, the claim that Atilla the Hun could have been two metres tall (actually, let us suppose, the tradition that he was quite short is correct). “Atilla the Hun” represents the person it does particularly, as well as singly (and incidentally should probably be thought to be a proper name rather than a definite description equivalent to “The Hun called Atilla”. Doubtless there were many Huns called Atilla — but only one was Atilla the Hun). According to the linguistic ersatzer or modal fictionalist, Atilla the Hun could have been two metres tall only if the world-book represents that, in some universe or other, Atilla the Hun (or his counterpart) is two metres tall. The world-book without the specification of the actual world apparently does not represent this: for it is uncertain which variable ought to be associated with Atilla, which with the property of being two metres tall, and so on, and so it is uncertain whether or not it represents Atilla as being two metres tall in any universe. This uncertainty is not merely epistemic, of course: the point is that it seems indeterminate or false that it represents that one of the universes as containing a two-metre tall Atilla the Hun (or a two metre tall counterpart of Atilla).

However, once it is specified that one of the universes is the actual universe (and it is as it actually is), then there is a guarantee that at least one of the universes contains the famous leader of the Huns (and the property of being two-metres tall). Given other principles which any adequate representational abstractionist theory will contain, it will follow that the object which is Atilla will be represented to be in other universes (or counterparts of him will), and that the height properties will be distributed in such a way so that the property which is Being Two Metres Tall will be instantiated by the object which is Atilla in some universe. The relating of the rest of the world-book to the actual world gives the world-book representations with particular content it would not have possessed otherwise, and since many of our modal claims contain particular content, if the linguistic ersatzer or modal fictionalist wants to tie the truth of our modal claims to truths about what is represented by the world-book in a perspicuous and reasonably standard way (necessarily \( q \) iff the world-book represents that for all universes \( q \), for example), the connection to the actual world is vital. This sort of connection is lacking with respect to the alien universals, of course, and is one of the reasons why representations involving them will lack the same capacity for particularity that representations of actual things usually have. Perhaps this lack of representative power with respect to universals not found in the actual world is the source of the confusion which leads to the conclusion that linguistic ersatzers or fictionalists cannot represent alien universals at all.
Thus far I have been to some extent avoiding an interesting issue that makes quite a difference to exactly what my proposed solution to the problem of representing alien universals provides for the linguistic ersatzer or fictionalist. As we have seen, linguistic ersatzers and modal fictionalists can allow that the world-book provides specific representations of alien universals, though not particular representations of them. To draw out the analogy between the power of one who can particularly represent and one who can merely represent, I used the example of *de se* content: the fact that I can make a representation with the sentence “I am in the Coombs building” which no-one else can make in the same way, either with that sentence or sentences like “Daniel Nolan is in the Coombs building”. The issue to be resolved is whether or not particular representation can be used to represent states of affairs which mere specific representation cannot, or whether they are merely provide different ways of representing the same states of affairs. This is a tricky question, and one which I am unsure about. The question is, I imagine, related to the attitude one should take in general to information gained about states of affairs through different “modes of presentation”, if I may use that expression in a somewhat non-technical sense.

In the classic case of *de se* representation, what to say seems unclear. It certainly seems that in at least one sense the sentence “I am in the Coombs building”, in my mouth, says the same thing as “Daniel Nolan is in the Coombs building”, in my mouth or anyone else’s. But famously it is controversial to say whether there is an important sense in which they say different things, and if so what the difference consists in. Lewis (1983) goes as far as saying that different propositions are expressed by the two claims, but others might wish to say that there is the one proposition expressed in two different ways, perhaps not even ways that are *a priori* knowable to be equivalent by anyone: after all, it seems that, at least in one natural idiom employing the word “know”, I could know that I am in the Coombs building without knowing that Daniel Nolan is, and vice versa. Suppose, in both cases, I had heard of the exploits of a person called “Daniel Nolan”, but had never realised that I was he. In the first case, I might know which building I was in but be unaware of where the much-talked of Daniel Nolan was lurking, and in the second might have become thoroughly lost and phoned a scrupulously honest and reliable friend for help only to be told, apparently unhelpfully, that Daniel Nolan was in the same sort of state I was in, and that he was in the Coombs building.

Another case where there is information gathered through different modes of presentation about the same state of affairs is what to say about Colour Mary’s first sight of a red object (see Jackson 1982 and 1986). When she has her first “red experience”, does she learn something new about red objects, or merely learn something she already new in a novel manner? This case is clouded by the metaphysical issue of qualia, but in terms of knowledge about the red objects at least (as opposed to Mary’s knowledge of
herself or the properties of her experience), the case is analogous enough so that the literature on Colour Mary provides a parallel.

The point can nicely be illustrated in the case of the world-book: when we specify that one of the worlds in the world-book is the actual world (and that the actual world is as it actually is) do we add a representation of something new, or do we merely allow the world-book to represent something that it already represented in a different manner? The added specification of the actual world will be helpful in either case, since in the second it provides the world-book with resources for particular representation which makes smoother the biconditional between what is true at worlds and what modal claims are true in the case of modal claims containing particular representations, whereas it is vital in the first case since otherwise the world-book fails to represent actually true propositions as being true at any world — and the necessary falsehood of actual truths is a burden no adequate theory of modality can bear. I am inclined to say that the addition of the particular representation of the actual world does provide representations of what is not represented by the world-book without that addition, so if I follow my inclination I am forced to say that we cannot represent aliens quite as well as we can represent at least some non-aliens, and that the Ug-worlders would be capable of making representations with contents we cannot capture. So the Ug-worlder’s claim “Ug is Instantiated” would capture something not actually represented by the world-book (since it is a particular representation unable to be made in the actual world), and the sentence in the Ug worlder’s mouth would express a truth. It is a possible representation which is true at a universe, but no representation of the world-book is an equivalent representation of the truth of that claim in a universe containing Ug and Ug-worlders.

Again, the analogy with representations having de se content might be invoked to show that this is not the sort of lack of representational capacity which should worry us. No representation produced by anyone else can represent in the same way what I can represent with the first person singular pronoun. But we nevertheless think that representations of this world produced by people other than myself can be perfectly adequate, and can potentially capture everything which we need to capture for most purposes. So it might be with representations of other possible universes: our lack of ability to make certain sorts of representations of universes containing alien universals does not mean that our representations cannot be useful for the purposes to which we wish to put possible worlds. The ersatzer and fictionalist may have to admit that there are possible true claims the content of which are not represented by the world-book, but this is an admission that can be made consonant with the overall picture: if the ersatzer thinks that the same propositions are being represented but it is merely the manner of presentation which is different, the ersatzer can point out that even though the representation which constitutes a particular claim about an alien cannot be a part of the world-book, still the world-book represents the proposition which would be expressed:
the world-book just represents it a different way. If, on the other hand, the ersatzer took
the stronger view that they are indeed different propositions which are represented by the
Ug-worlders when they speak of Ug, then the ersatzer must make a modification to her
theory. Instead of claiming a proposition is possible only if it is represented by the
world-book, the ersatzer should claim instead that this only holds for a certain class of
propositions (the general ones, or the ones capable of expression without particular
representation devices). A slightly more complicated story should be told for the
propositions which can only be expressed by means of particular devices: one approach
might be to appeal to the correspondence such propositions have to propositions
expressible generally (analogous to the correspondence between the propositions
expressed by “I am in the Coombs building” as uttered by me now, and “Daniel Nolan is
in the Coombs building”, as uttered by me now, if it were thought that claims with *de se*
content expressed different propositions from those which did not have *de se* content, as
for instance Lewis 1979 claims). An ersatzer or fictionalist could then say that a
proposition which could only be expressed employing particular representation is
possible only if the general proposition which corresponds to it is represented by the
world-book. This complication is perhaps an additional cost for the linguistic ersatzer or
modal fictionalist wedded to the idea of different propositions being expressed by
particular reference devices: but it is not a cost which cannot be borne.

7. From Universes to Worlds — Salvaging Ersatz Worlds From the Inter-
definition

Using only representative capacities which everyday language has the world-book has
managed to represent non-existing universes which correspond one-to-one with the
possible worlds. In fact, the world-book has only needed to employ the representative
powers of sentences employing only logical vocabulary (plus perhaps an assignment of
non-logical vocabulary to particularly represent actual objects — see below). But the
Ersatzer does not yet have an ersatz theory of possible worlds worth the name. For
ersatzers take possible worlds to exist (albeit they think worlds are *abstracta*), and the
linguistic ersatzer I have described has so far not provided existing candidates for the job
— for the universes his world-book talks about so far do not, by and large certainly,
exist. The modal fictionalist, on the other hand, might rest content with the universes
represented as the worlds — after all, they are only committed (*prima facie*) to the
existence of a modal fiction, and the world-book can serve as that. The ersatzer must do
more, or the method of providing specific representations of alien universals will not help
the linguistic ersatzer. Fortunately, more can easily be done, though the details are
fiddly.

Exactly what the details are will depend somewhat on the details of how the ersatzer’s
world-book is set up. Let me outline some of the alternatives, and proceed in the light of
them. Let me stress, however, that the outline of alternatives is by no means meant to be comprehensive, and there will be many matters of detail that I shall leave open. The question of what alternatives there are for the world-book is closely related to the question of how many alternative formal representations of possible-worlds systems there are: and it would be neither useful nor convenient to recapitulate every formal proposal (or possible proposal) in this area. I will just canvas some example ways of proceeding, but much is optional (particularly the choice of symbols to represent the different things which must be represented, of course).

I imagine that the initial inter-definition section will be organized so that it will have the syntactic features of predicate calculus. As well as the quantifiers and variables of standard first-order logic, it should also include a dyadic identity predicate, and logical devices to quantify over universals and express claims about their instantiations by particular objects. This can be done in many ways. Two of the most simple are i) introduce set-theoretic devices, including in particular a dyadic membership relation, and a device for representing ordered n-tuples of objects to represent instantiation of relations (a relation is instantiated by some objects if those objects form an ordered n-tuple for suitable n which is a member of that relation), or ii) introduce devices to represent instantiation other than set-theoretic devices. This can either be a dyadic Instantiation predicate, which functions formally like the membership relation (along with some device to simulate ordered n-tuples — take your pick), or there can be a different Instantiation predicate — each instantiation predicate being one -adicity higher than the universals for which it is used to represent their instantiation. So, for instance, their would be a dyadic Instantiation predicate to represent the Instantiation of monadic universals (i.e. properties), a triadic Instantiation predicate to represent dyadic universals (if a stood in relation r to b, then this might be represented by “Irab” (or its abstract equivalent), where that “I” represents the triadic instantiation predicate, for example). Subscribing the Instantiation predicates where they are taken to be a family of different adicities will be the convention I will adopt to distinguish them (so, for the example just given, “I,rab”).

For convenience, one might also introduce more than one type of variable. Different variables for universals and mere particulars might be convenient (e.g. capital X subscripted by ordinals to represent these universals: $X_1, X_2, X_n, X_{n+1}, X_n$, etc. and lower case x’s for individual variables). One could also keep track of -adicity of universals by means of an additional subscript or superscript. One also needs to quantify over universes, and an additional type (w subscripted with ordinals) is convenient to mark the distinction from other entities. Otherwise a “world-predicate” W needs to be introduced — the easiest way is to incorporate it as part of the “logical vocabulary” of the language of the representation. Whether or not more than one sort of variable is employed is not vitally significant, since anything that we want to say could be said either way.
What properties an object has varies from universe to universe (as does which relations an object stands in), and this needs to be represented by the world-book. There are two rival ways of doing this: quantifying over world-bound individuals which stand in counterpart relations to each other, or introducing world-relative predication. The first method is formally elegant, because statements about which properties and relations an object has can be made once and for all. The only resources that need to be added are a predicate to express an object’s being “in” a universe, and a predicate to express the holding of counterpart relations (on this formalisation “=” is reserved for intra-world identity only, while inter-world identity (if there be such) is analysed in terms of a counterpart relation holding between objects in different worlds whose pair are not in the extension of “=”). The usual dyadic predicate to express an object’s being “in” a world is a dyadic predicate “I” — but as I have reserved this for instantiation, a dyadic “B” (‘belongs to’) can be used to represent this predicate instead. Alternatively, being “in” plausibly reduces to the special case of the mereological part-whole relation (in systems with world (or universe)-bound individuals, at least) where the whole is a universe, and as mereological devices are likely to be introduced anyway into such a representation (though I will not specifically discuss these devices), the “in” relation may not require any new introduction of vocabulary. In principle the same could be said of the counterpart relation (or relations). If they reduce to relevant similarity, but for now I will assume that they are dealt with employing an unanalysed predicate (or predicates, if there are more than one counterpart relation). I bracket off the disputes about what constraints should be placed on the counterpart relation: the linguistic ersatzer can cut his cloth to suit his fashion in this respect.

Some have objected to the “world-bound-object” counterpart-theoretic conception of possible objects on the grounds that the denial of trans-world identity is implausible. But the formal machinery does not settle that issue either way, it seems to me: it is only when such formalism is given a certain interpretation that denial of literal trans-world identity, or indeed any specific meaning follows: the counterpart theoretic representation suggests that there is no trans-world identity, just as some models in physics suggest that three-dimensionalism about space and time is false. But a representation of the form outlined can be employed by those who think there is genuine transworld identity, providing suitable ancillary interpretation rules are in place (“=” may not mean identity simpliciter, for example: and perhaps one of the dyadic predicates which is formally a counterpart predicate is to be interpreted as the identity predicate, for example). Again, this is like the case of some models of spacetime employed in physics: three-dimensionalists, and even presentists, can (or at least they claim that they can) interpret these models so that they can be fruitfully employed by a good three-dimensionalist. However, producing a representation that will suffice does not rely on employing something analogous to the (formally) counterpart-theoretic representation. The
representation can also be interpreted as employing the device of trans-world identity between objects (or an object, more strictly speaking) found in different universes.

If an object is to be in different universes, and have different properties or stand in different relations in such universes, we require some device so that ascription of properties or relations to an object or objects can be relativised to a world. Exactly how this relativisation to a world should be understood is a matter that I will not and need not go into, but formally a convenient way to represent this is add another argument place to predications which are world-relative, an argument place which takes variables or constants representing worlds as its arguments (though at present we do not have constants in the world-book). If we accept for the moment a strong version of the necessity of identity (so that the identity predicate does not have world-dependent application), then the main predicate in the vocabulary I have specified for the world-book will be the Instantiation or set-membership relations. Instead of the Instantiation predicates I outlined earlier, the world-book would use devices analogous to those predicates but with each having an extra argument place, which would be occupied by variables representing worlds. So far this is all straightforward. However, once transworld identity is allowed for universes might need to be distinguished from maximal objects (i.e. the fusion of all the objects which exist at a given universe). Remember I am using “universe” in not quite its ordinary usage, but as a technical term for the first (concrete) conception of worlds outlined in chapter one. One universe can never be part of another. But some objects which are possibly maximal objects are not necessarily maximal. Take the fusion of everything that exists (our “actual world” in one sense of that phrase) as an example. As it happens, it is a maximal object. But it could be that everything that exists existed in exactly the way it does, save that there is a wormhole which leads to an otherwise isolated region which contains, as well as some spacetime, a hydrogen atom. If that were so, then plausibly the thing which is actually a maximal object would only be a proper part of the thing that would be the maximal object (the actually maximal object would not overlap the hydrogen atom in the case we are considering, for example). So we get a distinction between universes and possibly maximal objects, since the actual universe is necessarily not a proper part of any other

71 Set-membership might be thought to be modally rigid i.e. the extension of a set cannot vary from world to world. This is of course in conflict with a system which takes universals to be sets, accepts literal transworld-identity, and accepts that which universals are had by which object (or collections of objects) varies from world to world. The possible repairs are legion, and for the most part obvious. I will not go into them here.

72 Let me assume for the moment that the branching conception of “worlds” is not being employed. Not that a branching conception automatically delivers the consequence that some worlds are parts of others, but some might. The point I wish to make does not depend on the rejection even of this variety of conception of worlds, but the exposition of the point will be more tortuous than is necessary if I put it in a form of sufficient generality.
universe, while our actually maximal object can be a proper part of a maximal object. This line of reasoning can be resisted: one might wish to insist that in the possible case I described, it is the actually maximal object which is the maximal object (including extra spacetime and hydrogen atom), rather than its being identical to the duplicate of it which has all of its proper parts identical to it. I take this to be wildly implausible, and in any case it only pushes down the bump in the theoretical rug to have it appear elsewhere. For example, we cannot use the idiom of possible worlds straightforwardly to describe the case I have in mind, for a straightforward carrying across of the “in some universe” to “at some world” would give us that at the very same possible world a given hydrogen atom both exists (because of the truth of the possibility claim) and does not exist (because of the actual truth). Contradictions being too high a cost to pay, the theorist is left with a problem. Simpler by far to take steps to properly distinguish universes and maximal possibilia than to insist on their identity.73

The easiest way to do this is to resist the identification of universes with maximal objects, and to take them to be some other entity quantified over in the representation. The “belongs to” predicate then should not express part-whole relation, unless some metaphysics of constitution is invoked so that two things can have exactly the same members, arranged in the same ways, but nevertheless be distinct. No specially variable need to be introduced for the possibly maximal objects — they are possibilia just like the others, and can be adequately treated as such.

So much for a sketch of what kind of resources the representation might be interpreted as employing. From this base, it is relatively simple to define interpretations on further representations so that each corresponds to (or indeed, as the linguistic ersatzer argues, is identical to) a possible world. One group of candidates consists of the universe variables (or variables which are in the argument places of the “W” predicate, if there is no specific class of world-variables) — there will be at least one for each world, and only one for each world unless the representation also contains identity predicates with more than one variable as their arguments. (If so, then the candidates are the equivalence classes of world-variables where the equivalence is appearing as-an-argument-of-a-(non-negated) identity-predicate where the classes are not one-membered, or for any world-variables which do not appear in non-negated identity claims, then the relevant equivalence class is just the singleton of that world-variable.) These candidates will fill the theoretical role of possible worlds with a suitably defined “true at” relation, but they are not very much like the traditional linguistic ersatz conception of possible worlds, and

---

73 Provided we employ the mechanism of literal trans-world identity. Those who employ counterpart theory, on the other hand, do not need to worry as much, since maximal objects then strictly speaking cannot be part of more than one universe (though naturally they can have non-maximal counterparts). Those who employ the formalism of counterpart theory but believe in trans-world identity must provide a story to clarify the connection between universes and maximal objects, but there is no reason to suppose
are perhaps not terribly plausible candidates (though they are not too gruesomely
gerrymandered so as to be ruled out as potential candidates either, at least in my
judgement).

A more traditional class of candidates would be individual “world-stories”—a
host of representations, each of which represents the goings-on in one particular possible
universe. These can be appropriately defined given the resources of the world-book as I
have so far conceived of it, and would be the natural thing for a traditional linguistic
ersatzer to identify with possible worlds (they are like Robert Adams’s “world-stories”,
for example (Adams 1974). Again, there are probably many ways of accounting for a
connection between world-stories and world-books, but one recipe would proceed along
the following sorts of lines (I’ll assume a world-book with several sorts of variables for
convenience).

First, introduce constants corresponding to each variable, as “dummy names”
The easiest way to do this is by adding identity statements in the scope of the quantifiers
at the front of the following sorts (or their analogue): \( x_0 = a_0 \ldots x_\alpha = a_\alpha \ldots, X_0 = A_0 \ldots X_\alpha = A_\alpha \ldots \) for each sort of variable (optionally: except the world-variables). Once this is
done, a world-story can be constructed corresponding to each “universe variable” or
“world variable” recursively. The world-story corresponding to \( w_\alpha \) says that a given
particular object \( a_\alpha \) exists iff the world-book (the main representation) says that \( Bw_\alpha x_\alpha \),
the world-story says that \( IA_\alpha a_\alpha \) iff the world-book says that \( IX_\alpha x_\alpha w_\alpha \) (if we are using
world-bound predication), and so on, for all the different kinds of statements which a
world-story might be expected to make (such as instantiation of relations, identity
statements or diversity statements made employing the constants, and so on). Each of
these world stories will not strictly speaking form part of the world-book, and may even
be inconsistent with it, if they contain totality clauses, which ordinarily they would.
These totality clauses would say all the existing objects were identical to one or other of
the objects which the world-story claimed existed, and thus many world-stories would
not be consistent with the world-book, which says that objects which cannot all be found
at every world, all nevertheless exist (totality clauses would need to be added separately
after the rest of the construction). The introduction of constants was mostly to provide
the necessary anaphoric link between the distinct world-book and world-stories—it
could be done without constants if suitable anaphoric devices were employed, I suppose.
Each of these world-stories provides a complete intra-universe description of objects,
universals and their arrangement, and the “true at a world” relation is quite perspicuous
—a proposition \( p \) is true at a given world iff that world has as a part a representation that
\( p \) (or its parts jointly entail that \( p \), if some representation is left implicit). The ersatzer has
world-stories just like they have tended to want them—the only addition is that they
need the inter-defined representation of universes as well. It is not quite the traditional

that this cannot be done.
picture, of the representation structure representing all of the possible universes simply
being the aggregate of the representation of each universe: but it is still recognizably the
same sort of theory. A Linguistic Ersatzer can accept my modification in good
conscience.

8. How A Linguistic Ersatzer Can Reply to Lewis’s Objection Based on
Alien Universals

This discussion should already have addressed Lewis’s objection to linguistic ersatzism
that it cannot allow for possible worlds in which there are instantiations of alien
universals, but it might be worth going through his example specifically, as it is the most
prominent attack on linguistic ersatzism based on worries about alien universals. Lewis’s
strategy is to ask us to consider a world which stands to us as we stand to worlds where
some alien universals are: to consider a world which neither has in it all of the actual
universals nor the means to build up all the actual universals through structural means.
Lewis says

Think of an ersatzer philosopher who lives in a simpler world than ours.
The protons and neutrons (if we may call them that) of his world are
indivisible particles. There are no quarks; and so the distinctive properties
of quarks, their so-called flavours and colours, are not instantiated by
anything at all in his simpler world. They are alien [i.e. alien in Lewis’s
relative sense] to that world. ...
So... the philosopher in the simpler world... [does not] have any words for
the missing properties.

In saying that the philosopher does not have any words for the missing
properties, I take Lewis to mean at this stage that the philosopher does not have any
names (in the strict sense) for the properties. For he continues:

What we can do, however, is to speak of them by quantification. The
ersatzer in the simple world can introduce a language in which he can say,
what is false in his world;
Protons are tripartite. And there are natural properties X, Y,
Z, of which each least part of a proton has exactly one, which
enter as follows into the laws of nature that govern the binding
together of protons and their least parts:-X–Y–Z–. And the
properties X, Y and Z are different from charge, different
from spin, . . . .
where the dashed part is an open sentence in which the variables occur free,
and the dotted part is completed with different clauses for all the natural
properties he is in a position to name. That is, he can introduce a language in
which to formulate existential quantifications over properties. Such a
quantification is false at his world, but it could be made true by suitably
behaved properties alien to his worlds; and, thanks to the last clause, it
could be made true only by there being alien properties. Such sentences can
be members of, and true according to, his linguistic ersatz worlds (Lewis,
Lewis's argument strategy is to show that the philosopher in the simpler world cannot represent our actual world as a possible world, and then to appeal to our intuition that we could easily be in the example philosopher's predicament with respect to other, even richer worlds than our own. Of course the transition to the second claim from the first has been disputed (as by Armstrong 1989 p 56-57), and admittedly alien universals can be rejected even in face of Lewis's example if we think that the actual world is somehow special. For Lewis's argument to even get off the ground, however, he needs to demonstrate that our simple-worlded friend cannot represent the actual world in his world-book. Lewis argues for this premise as follows:

When that unfortunate philosopher in his simpler world constructs Ramsified ersatz worlds using the limited resources at his disposal, every world is at least partially described by one of them. But I say that the worlds with the alien properties — I mean worlds with natural properties that are alien to him, for instance our world — are described incompletely. He has said what roles for properties are occupied, but he has not said — and he could not possibly say — which properties occupy which roles. Here we are, with names for properties that he cannot name. We can distinguish our world from one in which, say, one of the quark colours has traded places with one of the flavours. The two possibilities are isomorphic, yet different. There are more ways than one to make one of his Ramsey sentences, or one of his Ramsified ersatz worlds, come true. Therefore such an ersatz world does not describe any of the relevant possibilities completely. (p 162)

I admit that our simple-worlded cousins cannot employ particular representations of quark colours and flavours, for example. As Lewis says, the philosopher in the simplified world cannot employ names for quark colour or quark flavour (though there is no reason to suppose Lewis would object to the claim that the ersatzer might be able to introduce dummy names corresponding to the variables in the quantified sentences). And maybe this means that the person in the simplified world cannot represent the actual world as comprehensively as we can, because it might be that to represent an object comprehensively essentially involves particular representation of it. But the claim that the ersatzer "does not describe any of the relevant possibilities completely" can be seen to be false if the ersatzer employs the suggestion I have offered, and provided complete description only requires specific representation of the thing described rather than particular representation of it. For the ersatzer that employs the world-book I offer her does not just say what roles are filled by properties in a world where protons are tripartite (and in other respects has the structural features which our world has). That ersatzer, even in a simpler world, distinguishes any given world with tripartite protons from any other where the natural properties of those parts are swapped around. The ersatzer in question can also distinguish possibilities which are "isomorphic yet different", and do so in the expected way — by also representing their differences. The ersatzer in the simpler world can do well enough to represent the sort of possibility which we take to be actual,
and possibilities isomorphic to it (albeit that s/he cannot do this in an isolated way), so we should not worry about the possible worlds which are richer than our world, as ours is richer compared to the simpler world of Lewis’s example.

9. Conclusion

We saw that the worry I labelled as Armstrong's worry was not very strong in the context of most theories of modality, since those that do not require that everything possible be represented can rest content with showing that general representation of propositions which would be true only if alien universals existed was possible. Now it has been shown that even against Linguistic Ersatzism the argument based on representation that alien universals must be impossible fail. Linguistic ersatzism and modal fictionalism can deal with this problem through being slightly more sophisticated versions of these theories than the examples in the literature have been, and by being more holistic in their representations of the possible worlds as a whole than has been traditional. The defence, on this point at least, rests.
Chapter 6 — Recombination Unbound

This chapter will discuss a principle of recombination for possible worlds (roughly, a general principle stating what can possibly co-exist with what), and a limitation that David Lewis has argued should be, and must be, placed on such a principle. I will argue that there is no need to endorse Lewis's limitation, and that Lewis's theory of possible worlds would be better without this principle. However, the interest and import of this discussion is not confined to its impact on Lewisian theories. Many people, myself included, wish to reject Lewisian modal realism (the doctrine that possible worlds exist and are concrete worlds just like our own) while still believing that some good and useful sense can be made of talk about possible worlds and possible objects. How a principle of recombination should be formulated is of interest to anyone who wishes to give an account of possible worlds and their content, regardless of the ontological status and nature of such worlds. I will not be talking specifically about how this might be done for other theories of modality, because once the option available to Lewis is understood, it is relatively straightforward to work out the analogues for other theories. Most other theories will have, implicitly or explicitly, a principle of recombination which has the same effect as the principle of recombination in Lewis's theory: and for those which have slight variations, the details may be different but the main point of this chapter should usually carry across. I will be focussing on Lewis's theory in this chapter for simplicity and clarity, but the lessons drawn can be of general import.

1. The Principle of Recombination and Lewis's Restriction

David Lewis thinks that the Principle of Recombination for worlds needs to be limited. To the basic statement of the principle, which is that for any objects in any worlds, there exists a world that contains any number of duplicates of all of those objects, there needs to be a proviso added, according to Lewis: "size and shape permitting" (Lewis, 1986c, p 89). The wording of the principle, which I have placed in italics, is my own, since Lewis does not provide an explicit formulation of his principle. However, I take it to capture the sense of Lewis's principle (minus his proviso), which he expresses at one point (Lewis, 1986c, p 89) as follows:

"Not only two possible individuals, but any number, should admit of combination by means of coexisting duplicates. Indeed, the number might be infinite."

The principle of recombination is meant to capture a plausible intuition we have about what kinds of things are possible. It is one way of expressing the idea that any sort of thing could possibly co-exist with any other sort of thing, and further that any sort of
thing could co-exist with many other instances of the same sort. This intuition is expressed in terms of duplication to avoid worries about "sorts" individuated by means of external properties or by external relation. For example, a thing of the sort "sole divine being" could not co-exist with others of its sort, for if there are two divine beings then neither is a sole divine being: but this is not a problem for the principle of recombination because it is restricted to intrinsic properties by means of the talk about duplication. Its expression in terms of duplication also enables Lewis to avoid being committed to transworld identities simply in virtue of the principle, as he would if he needed to allow that an intuition like "anything could co-exist with anything else" had to be cashed out by saying that for any two objects there was a world that contained both of those objects.

The proviso "if spelled out, would have to put some restriction on the possible size of spacetime" (Lewis, 1986c, p 103). Lewis presents two justifications for this proviso. The first is that if the unrestricted principle of recombination implies that there are possible sizes of space-time larger than we might have expected (for instance, there will need to be worlds which can contain more than continuum-many distinct things) (Lewis, 1986c, p 89). Since Lewis does not wish to have his recombination principle about objects deliver conclusions about the possible sizes of worlds, he thinks the proviso is needed. The second justification he offers is a stronger one. It is that a certain argument shows that the unrestricted principle of recombination leads to unacceptable results — even contradiction, according to Lewis (1986c, p 101). This is a stronger justification for the proviso because the former justification argued in favour of agnosticism about the unrestricted principle of recombination, while the latter demands its rejection. I will argue that the arguments relied on for the second justification in fact do not work, and so the unrestricted principle does not have the unacceptable consequences which Lewis claims it has.

The argument that is meant to show that the unrestricted principle of recombination leads to trouble is essentially one put forward by Peter Forrest and D.M. Armstrong (Forrest and Armstrong, hereafter F/A, 1984). This argument proceeds from an unrestricted principle of recombination to the conclusion that there is no aggregate of all worlds and no set of all worlds (F/A, 1984). Lewis takes this argument to be a reductio of the unrestricted principle, as the Forrest/Armstrong conclusion is unacceptable (Lewis, 1986c, p 92). I am happy to accept that if the Forrest/Armstrong argument showed that there was no aggregate of worlds, or if it showed that there was no class of possible worlds, that would be a reductio of the principle by Lewisian modal realist lights. I will show that it establishes no such conclusion, however.

---

74 Lewis also (1986c, p 91) talks of the Forrest/Armstrong argument as generating a contradiction, but as I read it his talk of it being contradictory implicitly relies on his assumption that the Forrest/Armstrong conclusion is unacceptable. If this is not being relied on, then in my opinion Lewis is just mistaken to think the conclusion of the argument is contradictory.
2. The Forrest/Armstrong Argument

Let us start with the Forrest/Armstrong argument, and see what conclusions can be drawn from it. A reconstruction which is faithful to the essence of the argument but which for ease of exposition differs in detail is the following:

1. Part of the Lewisian conception of possible worlds is that worlds have no parts in common (with a proviso about immanent universals, if there are any).

2. The principle of recombination is that for any objects in any worlds, there exists a world that contains any number of duplicates of all of those objects. So, in particular, it says that for any number of worlds, \((W_1, W_2, \text{etc.})\) there is a world such that it contains duplicates of all of those worlds, and that those duplicates are wholly distinct.

To argue for (2), we need to note two things. The first is that it assumes that worlds are in themselves: but given Lewis’s analysis of the “in a world” relation as the “part of a world” relation, this is not a problem (nor is it a problem for many who reject the “‘in’ as ‘part of’” analysis — on many people’s views, worlds are in themselves). The second is that instead of concluding that there is a world that contains duplicates of \(W_1, W_2, \text{etc.}\), it concludes that there are wholly distinct duplicates of those worlds in a world. This is more serious, as in fact the conclusion does not follow from the premise. To be fair to Forrest and Armstrong, they do not claim that it follows directly from the principle of recombination I have mentioned, and in fact try to motivate it independently on intuitive grounds. They are grounds that Lewis does not accept however (see F/A 1984, p 166). For now, let grant for the sake of the argument a stronger principle of recombination which will allow the Forrest/Armstrong argument to get started. Let us grant this Stronger Principle of Recombination:

\[
\text{SPR: for any objects in any worlds, there exists a world that contains any number of distinct duplicates of all of those objects}
\]

SPR will provide us with the premise we need (on the assumption that worlds are in themselves), for the Forrest/Armstrong premise is just a special case of SPR (where there is at least one distinct duplicate of each object, and all the objects are worlds). It follows from SPR that there is a world such that it contains distinct duplicates of all the worlds. Forrest and Armstrong draw the perhaps weaker conclusion that for any aggregate of worlds there is a world which contains distinct duplicates of all the worlds in
the aggregate (F/A 1984 p 165) from their premise, so for now I will stick with this latter
while I am laying out their argument.

Their argument continues as follows: Suppose, for purposes of reductio, that
there is an aggregate of all worlds (call it A). Then it follows, from the fact that for any
aggregate of worlds there is a world that contains distinct duplicates of all the worlds in
the aggregate (established above), that there is a world that contains distinct duplicates of
all the worlds in A. Call this world Giganto\textsuperscript{75}. Forrest and Armstrong then ask us to
consider the number of electrons in Giganto. Their argument then proceeds as follows:

Suppose $W_i$ is a world which is a part of A, and $W_i$ has just
N electrons. Then there will be some property F-ness (it may
be a relational one) which each electron in $W_i$ may or may not
have, and may or may not have independently of whether the
other electrons in $W_i$ have it. For each sub-set of the N
electrons it will be possible that precisely the electrons in that
sub-set have the property F-ness (F/A, 1984, p 165)

From this, we can work out that there are $2^n$ worlds containing electrons, since the power
set of a given set (that is, the set of all the subsets of that given set) always has a
cardinality equal to $2^g$, where g is the cardinality of the given set. Since Giganto contains
distinct duplicates of all of these worlds, Giganto must contain at least $2^n$ electrons. So
Giganto cannot be equal to $W_i$, and since the argument applies equally for any world we
choose, Giganto cannot be in A. But A was meant to be the aggregate of all the worlds.
By reductio, then, there is no aggregate of all the worlds.

There are some problems with this argument. One is its invocation of F-ness.
Nothing in any principle of recombination mentioned so far allows one to just assume that
there is a property F-ness that behaves in the way Forrest and Armstrong wish. Even
granting them their claim about F-ness, though, the argument still does not work. One
reason is that they talk as if there is trans-world identity of electrons. Not only is this
unfortunate because Lewis does not think that there is any such thing, and they are
supposed to be discussing a problem for Lewis’s theory, it is also unacceptable because if
it really was the case that there were the same electrons in each world, then the number of
electrons altogether would just be N, albeit that those N electrons would feature in many
worlds. Literal trans-world identity might also be thought to be inconsistent with the
premise, accepted for the sake of argument by Forrest and Armstrong, that each possible
world is entirely distinct from all of the others. We should therefore recast this argument
in terms of duplication. However, when we do so the argument looks less plausible. We
may grant that for any subset of electrons in $W_i$ there is a world with as many electrons in

\textsuperscript{75}They call this world $W_b$, but “$W_b$” is a mouthful, and isn’t as interesting a name as Giganto anyway.
it as there are in that subset, or even that for any subset S, there is a world $W^*$ containing as many electrons with the extrinsic property F-ness as there are members of S, and in addition $W^*$ contains as many electrons lacking F-ness as there are electrons in $W_1$ that are not in S. This, I take it, is the equivalent of their claim cashed out in terms of duplication rather than trans-world identity. Let us grant all of this. Does it follow that Giganto must be distinct from $W_1$?

It does not. For it does not yet ensure even that there are $2^N$ worlds with electrons in them. This is because we do not yet have any guarantee that the duplicate corresponding to one subset will not be identical with duplicates corresponding to other subsets. Remember that the condition that must be satisfied for all subsets of the electrons in $W_1$ is that there is a world which contains as many electrons with F-ness as there are electrons in that subset, and, if you like, as many electrons without F-ness as there are electrons in $W_1$ that are not in that subset. But this only requires that there be one world with N electrons only one of which has F-ness, one world where only two have F-ness, one world where only three have F-ness, and so on. The world with only one F electron will satisfy the condition for all the one-membered sets of electrons in $W_1$, the world with only two will satisfy the condition for all 2-membered sets of electrons, and so on. But this will only insure at most that there are N worlds containing electrons, not $2^N$. And since adding N to N N number of times (that is, multiplying N by N) only yields N, where N is any infinite cardinal, this does not even show that there are more than N electrons in the entire pluriverse. This defect could be remedied of course if we were allowed to insist on there being something like a function from worlds with electrons to subsets of the set of electrons in $W_1$ such that there were as many worlds with only one electron with F-ness as there were one-membered sets of electrons that were subsets of the set of electrons in $W_1$, as many worlds with only 2 electrons as there were 2-membered sets of electrons which were subsets of the electrons in $W_1$, and so on. But the principle of recombination does not justify insistence on any such thing.

The Forrest-Armstrong argument, then, is not successful. Even if one grants the strengthened principle of recombination to them, and even allowing that there is an extrinsic property Fness which functions in the way they wish, their argument does not even secure them the conclusion that there is no such world as Giganto.

David Lewis, in his discussion of the Forrest/Armstrong argument, provides a recasting of the objection that does not rely on "F-ness" in the way that the original argument does (though it too must rely on SPR76). It faces its own specific problem however, and in

---

76 Lewis does not produce the stronger requirement that the duplicates be distinct when he introduces his principle of recombination in section 1.8 of (Lewis, 1986c). However, he does claim that a principle that amounts to SPR is the principle he endorses in 1.8 when he talks about "copying" on p 101 of (Lewis, 1986c). I take it that even if this is not tantamount to endorsing SPR, it is at least good evidence that
addition suffers from the same basic flaw. Let us look at this version briefly. Lewis's recasting of the Forrest/Armstrong argument goes as follows:

Suppose the big world [i.e. Giganto] has \( K \) electrons in it. Then there are \( 2^K - 1 \) non-empty subsets of the electrons of the big world; and for each subset, there is a world rather like the big world in which just those electrons remain and the rest have been deleted. (I [i.e. Lewis] take this to be a subsidiary appeal to recombination.) Call these worlds variants of the big world...

There are \( 2^K - 1 \) variants; there are non-overlapping duplicates of all these variants within the big world; each variant contains at least one electron, therefore so does each duplicate of a variant; so we have at least \( 2^K - 1 \) electrons in the big world; but \textit{ex hypothesi} we had only \( K \) electrons in the big world and \( 2^K - 1 \) must exceed \( K \); so the big world has more electrons in it than it has. For a world to be bigger than itself in that sense is indeed impossible; which completes the \textit{reductio}. (Lewis, 1986c, p102)

Note that when Lewis says that "there is a world rather like the big world in which just those electrons remain and the rest have been deleted" he presumably intends that to be cashed out not in terms of literal identity of electrons between worlds but in terms of duplicates. After all, Lewis famously does not believe in literal trans-world identity of things like electrons, and his principle of recombination is meant to be about duplicates of objects in different worlds, not the reoccurrence of the objects themselves.

The problem specific to Lewis's reformulation of the argument though is that the "subsidiary appeal to recombination" just does not serve to produce the desired result. The principle appealed to, namely that \textit{for any objects in any worlds, there exists a world that contains any number of duplicates of all of those objects} does not allow us to say that for a given subset of electrons in a world, there exists a world with \textit{only} as many electrons as there are in the subset. At least, it does not when the principle is being read as being equivalent to the following:

\[
\textit{for any objects in any worlds, there exists a world that contains, for any number, at least that number of duplicates of all of those objects.}
\]

Suppose, however, that one wanted to read it as the following:

\[
\textit{for any objects in any worlds, there exists a world that contains, for any number, exactly that number of duplicates of all of those objects.}
\]

Lewis would not have any trouble accepting SPR (apart from wanting to add the \textit{size and shape} proviso, of course).
This reading would be untenable, as it would lead to a contradiction. After all, consider two objects A and B, such that B is composed of two duplicates of A. The current principle tells us that there is a world which contains exactly one duplicate of A and exactly one duplicate of B, which is not possible. It is better to read the principle of recombination in the first way rather than the second, if one is to be fair to Lewis's intention. This flaw in the argument could be repaired by invoking a more complicated principle of recombination, but to state such a principle is not a straightforward matter.

However, the more important flaw in Lewis's recasting is the problem that I pointed out for the original Forrest/Armstrong argument. Lewis's argument is not good enough, for all that it shows is that there is a world with only one electron, a world with only two, one with only three...and so on, so that for every size of a subset of K there would be a world that contains exactly that many electrons. What it does not show is that there is a distinct world containing electrons for every subset of K. That is, there could be, for all that this principle gives us, only one world with one electron, only one world with two... up to there being only one world with K electrons, rather than there being as many one-electron worlds as there are subsets of K containing only one electron, as many two-electron worlds as there are 2-membered subsets of K, and so on up to there being as many worlds with K electrons in them as there are K-sized subsets of K. The current reading still does not licence us to draw the conclusion that Giganto must have more electrons than it has. Lewis's reformulation does not deliver the desired (or as the case might be, undesired) conclusion that there is a world bigger than it is. Let us leave the Forrest/Armstrong argument and Lewis's recasting then, as it seems that these arguments will not serve to provide the motivation for Lewis's "size and shape" restrictions.

While the Forrest/Armstrong argument and Lewis's recasting are not satisfactory for reaching their desired conclusions, a better argument relying on the unrestricted principle of recombination can be constructed to deliver similar conclusions to those desired for the Forrest/Armstrong argument. Let us therefore look at this argument and examine whether its conclusion justifies Lewis's restriction on the principle of recombination.

There is a quick and simple argument from the principle of recombination (For any objects in any worlds, there is a world that contains any number of duplicates of all of those objects) to the conclusion that there can be no set of all possible objects — that is, there cannot be a set of all the objects to be found in worlds. The argument is as follows: suppose (for reductio) that there is a set of all possible objects. This set must have a cardinality, as it is part of the definition of cardinality that all sets have it — call it C. But if it has a cardinality, then there must be a greater cardinality than it (e.g. the cardinality of its power-set). Call one such cardinality C*. From the principle of
recombination, for some object, there is a world that contains $C^*$ duplicates of that object. So there are at least $C^*$ objects to be found in worlds, so the set of all possible objects must have at least $C^*$ members. But $C^*$ is of course strictly larger than $C$ — so the set of possible objects (with cardinality $C$) must be larger than itself. *Reducio.*

How bad is this conclusion? So far, not very. After all, there could still be a *proper class* of possible objects. Proper classes, as they are often conceived, are just like sets except that they are not themselves members of classes or sets. Exactly what one wants to say about the cardinality of proper classes may well be controversial. Some may wish to deny that some proper classes have cardinalities at all — perhaps because cardinality is only well-defined for sets, or the notion of cardinality involves the notion of functions and for some reason functions fail to apply in various ways to (at least some) proper classes, and so on. Even if the notion of cardinality of proper classes is acceptable, though, it is surely plausible that some proper classes are such that no class has a greater cardinality than any of them. Consider, for example, the class that is the union of all classes. It is hardly going to be the case that there will be a class of a higher cardinality, where this implies that there is no function from the second to the first. After all, the second will be a sub-class of the first, and a class cannot have a subclass of higher cardinality than itself! So there is no straightforward cardinality objection if the possible objects are taken to form a proper class. Furthermore, there is no argument available (so far as I can see) from the principle of recombination to the conclusion that there is no aggregate of all possible objects. Even if they did not form a set, why should one suppose that there is no aggregate of them? So it looks like it might be the case that if we want to accept the completely unrestricted principle of recombination, we need to give up the notion that there is a set of all possible objects, but we can retain the notion that there is a class of such objects, and also that there is an aggregate of such things.

The conclusion that was originally wanted — that there is no set or aggregate of possible *worlds* — does not seem to be derivable, however. If it could be shown, or if the principle of recombination insisted, that there was an onto function from the class of

77 See Bernays (1937-1954) or Goedel (1940) for two classic accounts of proper classes, and see Lewis (1991) for Lewis's account of proper classes. Two things that are worth noting with regard to the treatment of proper classes: the first is that I will be following those who take it that classes that are not proper classes are to be identified with sets, as opposed to the view that a set and a class that have the same extension are nevertheless distinct entities. Secondly, that proper classes are not uncontroversially part of set-theoretic orthodoxy — Zermelo-Frankel set theory does not mention such things. Von Neumann-Bernays-Goedel set theory has had its distinguished proponents, and a theory which includes proper classes is one which Lewis himself endorses, so at least as far as this is a chapter about why Lewis should or should not restrict his principle of recombination, an appeal to proper classes is in order.

78 If I may insert an *ad hominem* here: Lewis is already committed to there being objects which form an aggregate but not a set. Consider his objects which are mixed fusions of individuals and classes (Lewis, 1991,p 8). For any two of them their will be a fusion of those two (from unrestricted mereology), but neither is a member of any set (or class), and the fusion of the two will not be either.
possible objects to the class of possible worlds, or even an onto function from mutually
distinct possible objects to the class of possible worlds (which captures the intuitive idea
that “anything could fail to co-exist with anything else”), then one could use the fact that
the possible objects cannot form a set to show that there cannot be a set of possible
worlds. But the sort of principle of recombination Lewis would be interested in would
only be one that is committed to the existence of duplicates corresponding to possible
objects in other worlds — and one object can do duty as a duplicate for countless others.
(Exactly how the intuition that “anything could fail to co-exist with anything else” should
be accommodated by Lewis is a somewhat complicated question — one that I will not
bother to address here, but one which Lewis owes us an answer to).

The thought that the possible objects may not yield a set might be thought to be
bad enough. Lewis says, in fact, that “we have no notion of what could stop any class of
individuals... from comprising a set” (Lewis, 1986c, p 104), and there are other
problems mentioned by both Forrest/Armstrong and Lewis for preferring that the possible
objects form a set. In the next section of the chapter, therefore, I will examine the option
of treating the class of possible objects as a proper class, and will answer objections
which one might have to this strategy.

3. How *possibilia* can be seen as forming a proper class

So far, the only positive argument for saying that Lewis should allow the things in
possible worlds form a proper class has been the intuitiveness of the unrestricted principle
of recombination. Before I examine any other positive arguments for this conclusion, let
me first examine some of the arguments against it, based on considerations raised by
Forrest/Armstrong and Lewis.

The first type of argument is that we (or Lewis, at least) require sets of various
possible objects to serve various theoretical purposes, and that if there turned out not to
be such sets, then Lewis’s project would be in serious trouble elsewhere. Avoiding such
serious trouble might then be a sufficient reason to restrict the number of individuals in
worlds. Forrest and Armstrong present an argument which is more or less this one when
they are discussing the prospect that all the worlds and, *a fortiori*, all the things in them
do not form a set. They point out that Lewis’s account of properties is that they are sets
of objects. The property of *being an electron*, to use the Forrest/Armstrong example
again, will be the set of all electrons, no matter where in the pluriverse they be. If we
take there to be no set of electrons, then, what happens to Lewis’s conception of the
property of *being an electron*? Lewis makes a similar point (Lewis, 1986c, p 104)

79 An extremely similar point can be made about relations in Lewis’s system too, since for Lewis
relations are sets of ordered pairs (or ordered triples, or in general ordered n-tuples depending on the adicity
of the relation in question). My conclusions about properties will carry over to relations in an obvious
about his use of sets of possible objects and possible worlds to provide semantic values for words. The attraction of being able to do Lewisian semantics or Lewisian accounts of properties and relations is surely (for Lewis at least) worth imposing a minor restriction on his principle of recombination. Can he get the benefit of sets of possible objects even without restricting the principle of recombination?

I think he can. The first thing to note is that even when possible objects do not form a set, they still form a class — and classes are suited for nearly all the purposes that sets are. That there would be no set of electrons, but only a proper class of them need not mean that the Lewisian account of what it is to be an electron must be abandoned— for the property of *being an electron* will be the proper class of electrons rather than the set. Lewis does have a reason though, for thinking that proper classes will not be adequate substitutions for sets to do the work of properties. In a footnote (Lewis, 1986c, p 50 n37) Lewis points out that he wants properties and relations to themselves have properties and relations. To have a property is to be a member of that property for Lewis, and proper classes are not members of any other classes, so proper classes that were properties would not themselves be able to have any properties. Here then is a reason to deny that the electrons, for example, form a proper class — if they did, the property of being an electron would not itself have any properties, and Lewis wants to say that it does (or at least wants to leave it open that it does). The use of set-theoretic constructions out of semantic values is also important in Lewis’s semantics. Can this conflict be avoided, or does accepting the unrestricted principle mean that Lewis is denied an adequate account of properties and semantics?

It can be avoided, and in at least two ways. I will only discuss the ways of avoiding the problem of no longer having sets of the relevant possibilia in the case of properties, and will leave the discussion of how to save his formal semantics for some other time, both because the case of properties will be much easier to discuss than the intricacies of formal semantics, and because the solutions offered for properties can be extended to solve the problems raised for formal semantics in a relatively straightforward way.

The problem for properties, remember, is that we wish to allow that properties themselves can have properties and stand in relations. If properties are conceived of as being classes of their instances, then those properties which would have proper-class many members (like electrons, or indeed like any natural property) would not themselves be able to be members of any properties or relations, as proper classes, as we have been discussing them so far, are not members of any classes. The first solution to this problem is the sort that I would favour: to allow into one’s ontology sparse universals.

manner, and so I will not be discussing them separately.

80 Lewis’s use of the word “class” in this footnote is not the usage I have adopted (and which Lewis later
Lewis himself thought that this was about as good an option as his theory of "naturalness" of sets at one stage (Lewis, 1983, and Lewis 1986c, p 64), though he seems to have become less favourable towards the idea since (Lewis 1986a, p 26). How will it help here?

Here is one method: identify the properties and relations, (or at least those properties and relations which have some degree of naturalness) not with classes of instances, but with universals and constructions (perhaps set-theoretic constructions, perhaps not) out of them. Something will then be an instance of a property, not in virtue of being a member of that property, but in virtue of something else. This "something else" can either be a relation of Instantiation holding between the properties or relations and the individual object(s), or perhaps just the truth of some primitive predication using something like the two-place predicate "...has..." or "...instantiates..."81 if, like Lewis, you think that realism about universals is still committed to some primitive predication (Lewis 1983, p 353-354). The extensions of the Instantiation class for each universal will be proper classes, but the properties themselves can easily have properties or stand in relations — this is just for one universal (or set-theoretic construction out of universals) to be an instance of another. One still has the classes of instances available for any special uses one might prefer a class of instances for, but one has properties (and relations, for that matter) that can easily be themselves instances of other properties and relations, and members of classes of instances of such other properties and relations.

One could even, if one liked, reserve the names "property" and "relation" for instance classes rather than the universals. Then one could not say that properties with proper-class many members themselves were members of any other properties, but one could say that the universal (or construction out of universals) corresponding to each natural property or relation was an instance of other universals, which would do the same theoretical work — it is just a notational variant of the view I outlined on which we called the universals the properties and relations, after all.

I should mention two things quickly in order to deal with suspicions one might have that this project faces serious difficulties. The first is the matter of "constructions" out of universals — one might think that Lewis would not like such things, given his well known attacks on structural universals (see especially Lewis 1986a). These constructions need not be structural universals in Lewis's (and others') sense. As an example of what else they might be, they might be set-theoretic constructions out of universals — for instance, the "property of being blue or green" may just be the set of Blueness and Greenness, and to have the property might simply be to be an instance of something in the set. To be an instance of Methane (or Methaneness) might be to be related in the

adopted in (Lewis, 1991)), but the point he makes is still valid even with the different usage.

81 These are meant to be in quasi-quotes, of course — the fundamental facts of the universe don't rely on
appropriate way to a sequence having as members such things as Hydrogen, Carbon, Covalent Bonding etc. These examples are fairly schematic, but something along these lines might well do.

The second concern that might be expressed is that the introduction of sparse universals and constructions out of them to fulfil various duties demanded of the natural properties leaves the question of the unnatural properties untouched. Many, in fact most, of the classes of individuals will be hopelessly gerrymandered, and not natural to any significant extent. How are we to account for the properties and relations that such unnatural properties might be supposed to have if universals (and constructions therefrom) correspond only with the fairly natural? One solution to this problem would be to have some infinite and apparently fairly gerrymandered constructions out of universals to be the things corresponding to gerrymandered classes. But one need not do this — one might also reasonably deny that the completely unnatural classes do not have any properties or relations over and above those that trivially supervene on mereological or set-theoretic connections such classes have. The fact that the utterly gerrymandered properties that these classes are to be identified with have no significant properties or relations is something that one might be prepared to accept — since the only properties and relations that one’s theory will be concerned with, by and large, will be the somewhat natural ones, and I would guess the requirement that properties and relations themselves have properties and relations is only pressing for the fairly natural ones.

There is a second alternative available as well — one which is closer to Lewis’s current theory of properties, and which does not require ontological commitment to universals. So far I have been talking of proper classes as if their distinguishing feature is that they lack singletons — that is, that they are not members of any sets. But there are other ways of conceiving of the distinctive feature of proper classes. Another way of conceiving of them is that they are the classes as large as Reality: that is, they are the classes that have as many members as there are things (both individuals and classes) altogether. Lewis uses the technical term “Large” to describe the size of such classes. It is controversial even that all proper classes have this feature, but Lewis at least accepts that they do (see e.g. Lewis 1991, p 95, p 98). If this criterion of size is taken as the defining feature of proper classes, it may then be open to us to claim that in fact some of them do have singletons. (Not all of them can, on pain of various set-theoretic paradoxes). Again Lewis should not be too concerned by this — one of the weakenings of his axioms he says would be acceptable leaves it open whether or not some large classes have singletons. Since the only properties and relations which themselves need to have properties and relations which do not just trivially follow from their mereological or

the meanings of words in English.

82 The weakening of Domain to Domain+, discussed in (Lewis, 1991) at p 145.
set-theoretic features will be the ones which are significantly natural, we can furthermore say, if we wish, that it is all and only the natural proper classes that possess singletons.\(^{83}\) Lewis's theory can remain almost untouched — for proper classes with singletons can perform the roles Lewis desires for natural properties and semantic values. However, proper classes will still all be of the largest size, so the unrestricted principle of recombination will not drive us to have classes of possibilia larger than these natural proper classes. If this path rather than the path of universals is taken, however, a minor modification of Lewis's theory of classes should be introduced: His axiom Domain (see Lewis, 1991, p 95) should be rejected, in favour of Domain\(^{+}\) (see p 145) or something else which is stronger but which is inconsistent with Domain.\(^{84}\)

With the problem of properties and semantic values out of the way, let us consider other objections one might have to the unrestricted principle of recombination and the thought that possible individuals form a proper class. Apart from the point about properties and semantic values, the other consideration relevant to the present proposal that Lewis raises against the thought that the possible objects do not form a set goes as follows:

How could the worlds [or, presumably, the individuals] possibly fail to form a set? We do say that according to the iterative conception of sets, some classes are ‘too big to be sets’, but this is loose talk. Sheer size is not what matters; rather, the obstacle to sethood is that the members of the class are not yet all present at any rank of the iterative hierarchy. But all the individuals, no matter how many there may be, get in already on the ground floor. So, after all, we have no notion what could stop any class of individuals... from comprising a set (Lewis, 1986c, p 104)

Lewis's views on sets have developed since this remark, and he may well no longer wish to stand behind these comments. But in any case the argument here does not seem to me to be compelling. For a start, surely it is "loose talk" to talk of things being "not yet all

---

\(^{83}\) It is a reasonable assumption that the natural classes will be Few, in Lewis's technical sense (see Lewis, 1991, p 90). Even if they were Many, or even if the class of natural classes was Large, however, there would not necessarily be any difficulty.

\(^{84}\) My suggestion is "Domain' " to be defined as follows: any part of the null set has a singleton; any singleton has a singleton; any small fusion of singletons has a singleton; some large fusions of singletons have singletons; and nothing else has a singleton.

some alternatives might be

**Domain\(^{n}\):** Any part of the null set has a singleton; any singleton has a singleton; any small fusion of singletons has a singleton; any natural fusion of singletons has a singleton; and nothing else has a singleton.

**Domain\(^{+}\):** Any part of the null set has a singleton; any singleton has a singleton; any small fusion of singletons has a singleton and some (few) large fusions of singletons have singletons.
present at any rank of the set-theoretic hierarchy”. Set existence is not a temporal matter, and the talk of construction is only metaphor. On the other hand, I believe that perfectly literal sense can be made of the claim that some class is a proper class because it is ‘too big to be a set’. If there is a class of things such that there is no set whose members stand in one-to-one correspondence with the members of the class, then that class fails to meet a necessary condition for forming a set, and fails to meet this condition in virtue of being “too big”. How is this loose talk at all? Finally, it is difficult to argue against those who claim that no-one can have a certain notion — I can do little more here than insist that I, at least, seem to be able to entertain such a notion, and cannot see any reason why the option I am exploring is literally inconceivable.

There is a more serious worry here however: that the suggestion that the individuals form a proper class is out of step with the iterative notion many have of sets, and is out of step with mathematicians’ practice and intuitions. If my proposal was to fly in the face of the professional opinions of mathematicians, this might be a worry. In actual fact I think such a worry is not particularly great. For one thing, the orthodox theory of sets employed by mathematicians — pure set theories such as ZF, for example — do not deal with individuals at all, and opinion among prominent set theorists seems to vary from being silent on the question of individuals to adopting, for convenience, the thesis that there are no individuals at all. Proper classes of any sort do not appear in many orthodox treatments of sets, so the question of whether one sort of thing that is not even discussed do or do not form another sort of thing that is not even discussed is hardly settled by such theories. This is not to say that there is anything wrong with these sorts of theories of course — though I would tend to see pure set theory as only a fragment of the entire story about sets. The rhetoric of iterativeness is strong among contemporary set theorists, it is true — but much of an iterative conception can still be salvaged after my suggestion, and the iterativeness of ZF and other treatment of pure sets remains untouched. In any case, were mathematicians to rule decisively that the individuals did form a set, this would be to make a substantive claim about the number of individuals there are (or, perhaps more strictly, what proportion of Reality is composed of individuals). I am happy for mathematicians to tell us about mathematical ontology — but when it comes to the ontology of individuals, and questions like “how many individuals are there”, it seems to me that metaphysicians have at least as much a right to an opinion on these questions as pure mathematicians, and probably more right. Finally, the rhetoric of iterativeness is often not meant literally (it can be seen as being of one piece with the metaphor of temporal construction of sets), and when it is meant literally it is sometimes an expression of a constructivist philosophy of mathematics which is by no means uncontroversial or orthodoxy. So even if this did fly in the face of mathematical opinion and the rhetoric of iterativeness (and, as I say, I do not believe it does), this could
be attributed just as much to mathematicians judging questions outside their brief as to a metaphorician daring to hold a mathematical opinion.85

The objections to believing that there are a proper class of individuals do not seem overwhelming. But there is another worry that the Lewisian defender of the unrestricted principle of recombination must address if their theory is to be accepted. It is the worry Lewis expresses when he first addresses the question of what principle of recombination to adopt in *On the Plurality of Worlds* (Lewis, 1986c, p 89-90). In this section he argues that there should be the proviso “size and shape permitting” because we do not have good grounds for ruling out the claim that there is some maximum size of spacetime, and the unrestricted principle does rule that out. Lewis appears to think that ruling out a maximum size of spacetime is both unnecessary86 and undesirable.

The fact that the unrestricted principle of recombination is more intuitive than the restricted version is presumably *prima facie* evidence that we should decide in favour of the unrestricted principle, rather than accept a contrary or suspend judgement.87 But this *prima facie* reason may fail to be adequate support for the principle for two reasons. One might think that there are also positive things to be said for a contrary (specifically, that there is a maximum size of spacetime), or one might just think that the *prima facie* reason does not itself overcome the ‘inertia’ of the default position of suspended judgement. By defending the unrestricted principle of recombination from objections which might be raised to it, I have been *ipso facto* dealing with considerations which might be thought to be positive arguments for the thesis that there is a maximum size of worlds (hereafter “MS”). To argue directly against the case that one should suspend judgement on the restriction is harder. The next section of this chapter will examine some arguments against MS (besides the fact that it conflicts with the intuitively plausible unrestricted recombination principle): while I will claim that two arguments against MS that sound plausible are not satisfactory, the arguments I think more promising are hopefully adequate to justify the rejection of MS, rather than mere agnosticism about it.

85 This might sound like I think that there are firm inter-disciplinary boundaries, and demarcation disputes about who is entitled to have an opinion on what are worthwhile enterprises. I don’t in fact think any such thing — I make these remarks mainly for the benefit of those who think that there is something odd or improper about philosophers disagreeing with mathematicians about sets (or classes).

86 In the epistemic or doxastic sense.

87 I suspect this should be even more true for Lewis, who sees the project of metaphysics and philosophy generally as just to improve the intuitive theory we begin with (informed by our empirical discoveries, of course), rather than to provide ultimate grounds or to revise our theory of the world(s) wholesale. However, I will not go into this here.
I will begin with two arguments which, while having an initial attraction, are not really good arguments against MS. Part of my reason is that these were arguments that I found tempting at first (and perhaps am just reluctant to forget), but if further justification is needed, it is that others seem to have shared what I now take to be my errors: Forrest and Armstrong provide an argument similar to the first (F/A p 166 & p166-167n), and Lewis is worried by a consideration similar to the second (Lewis 1986c, p 103).

The first of these two reasons to suppose that there should be no maximum size of worlds is that models of worlds of any size can be provided. One way is that outlined in (F/A p 166-167, n1), but there are other obvious ways of constructing such models as well. Checking to see if there is a model of a putatively possible situation is a standard practice of modal reasoning, and the extent to which we are prepared to deny the possibility of something even though it can be modelled using set-theoretic modelling is an extent to which we lose one of the few epistemic methods we have (by Lewis’s lights especially) for resolving modal questions. (How we can know of set-theoretic models seems to be an equally difficult puzzle on Lewis’s view, but this is another matter). This cannot be a very weighty argument however, since of course models are not always suitable: where we have not imposed sufficient conditions (such as meaning postulates or other such) on predicate letters, for example, we can apparently have models where, for instance, all spinsters are married, or water is XYZ. One might wish to claim a similar thing here — the supposed models are misleading because the relation relating the objects in the model does not satisfy all the conditions which apply to the sort of relation Lewis thinks is necessary to unify worlds. Lewis holds that relations that unify worlds must be either "analogously spatio-temporal" (as he says in Lewis 1986c, p 75-76) or perhaps they only need to be unified by a natural external relation of the sort discussed at (Lewis, 1986c, pp 76-78). If such relations have to meet further conditions than the relations in the putative models of worlds of very large cardinality, this failure of the putative models to model modal reality is analogous to the sort of failure suffered by a model that does not respect the fact that (let us suppose) water is never XYZ.

One worry that the existence of models may give rise to (and this, rather than the objection from the mere existence of models, may be the force of the comments in F/A p 166) is that any particular size meant to be the maximum size of worlds would be arbitrary or ad hoc. Lewis seems worried by this (Lewis 1986c, p 103), and pins his hopes on mathematics delivering at least one size, and perhaps more, as sizes which look like a non-arbitrary place for the maximum to be. If Lewis had to pin his hopes to such a turn-up, then he may well be in trouble. We have, it seems to me, a reasonably good grasp of how the transfinite is laid out, and no natural break stands out particularly as an
obvious place to locate the maximum size of possible worlds.\(^{88}\) I do not see that he does, however, and I disagree with him that there would be something objectionable in assuming that the maximum size may well appear arbitrary. Of course, were his theory to state which size the maximum is meant to be, and that without any justification, then that would be objectionably arbitrary, but if the MS thesis is restricted to the claim that there is some maximum size, and that for all we know that size would seem arbitrary to us should we consider it, I do not see what theoretical failing the theory would have suffered. In fact, it seems to me that to not commit oneself to a given size as the maximum size but also insist that the maximum size will be one that would appear non-arbitrary to us itself seems objectionably arbitrary. Lewis also has a worry to do with the thought that the way things are in the modal universe cannot be arbitrary, for if they were it would suggest that the modal universe could have been otherwise, which is impossible (Lewis 1986c, p 103). I neither feel the intuitive pull of this thought nor can see what argument it might be elliptical for, so I cannot at present do it justice. As far as I can tell, then, the worries discussed so far do not count against MS.

Some objections to MS which I believe do have some bite can be seen once MS is cashed out in modal terms. The thesis that there is a maximum size to possible worlds entails that necessarily, space-time\(^{89}\) is of a certain size or smaller. How much stronger it is than this depends on what other things besides space-time and spatio-temporal relations can unify objects into worlds. As I mentioned, Lewis’s worlds can be unified by relations other than the spatio-temporal relations\(^{90}\) found in actuality. At the very least, they can probably be unified by relations which are “analogously spatio-temporal”, and perhaps they can be unified by other some other sorts of natural external relations. So MS may well commit one to the thesis that the extent of interrelation by any analogously spatio-temporal relation, or perhaps even natural external relation is necessarily limited.

To put it another way, it is the thesis that some relations are necessarily such that they have at most a given number of instances. This would be a strange sort of necessity, differing both from those sorts that seem very general (such as the necessities of logical

---

88 At least, none of which I am aware. I may of course be seriously mistaken about how good a grasp of the transfinite we have, or maybe my modal intuitions are not honed appropriately, but I have not come across too many people thinking that a certain transfinite cardinal is the obvious limit either.

89 I do not wish to make any assumptions here about the debate between relativists and absolutists about space-time. Relativists (or Relationists, as they are less misleadingly called) can presumably put some gloss on sentences containing the nouns like “space” and “time” so that they express truths in people’s mouths quite often — I invite Relationists to apply such a gloss to my talk of “space-time” — or at least to take my talk to not be committed either way as to whether the Relationist gloss is the perspicuous way of understanding it.

90 Those absolutists who see space-time as primarily substantial and not involving fundamental relations are also invited to apply their gloss to my speech, or at least take me to be leaving it open whether or not their gloss is to be applied.
relationships) and those which may be more specific but fall under categories which, even if mysterious, we at least have some grasp of. These might include (if there are such things) necessity arising from mereological connections or other internal relations, analyticity or reference constraints, and perhaps other categories. But a necessary maximum to how many instances of an external relation there could be seems mysterious even by the murky standards of metaphysical theories. Such mysterious necessity should be rejected, and to do so entails the rejection of MS.

Another unwelcome feature of MS (though not as serious) is that it not only deals with things one might want to think of as the “very large”, it also has consequences for something which might be intuitively thought of as the “very small” (though as far as the technical sense of the word goes (see Lewis 1991 p 89), each area is equally to do with the Large). MS would rule out, apparently a priori, the metaphysical possibility of a certain sort of “gunk”. “Gunk”, used in its technical mereological sense (see Lewis, 1991, p 20) is a word for a sort of thing which is such that any of its parts will themselves have proper parts. Gunk is not (metaphysically) atomistic, it is ‘infinitely divisible’. Even if there is no actual gunk, it is plausible that gunk is metaphysically possible — or at least, if someone is to take metaphysical atomism to be a necessary truth, then the onus is on that person to provide us with a reason why. MS, of course, can allow for the possibility of one sort of gunk: for a piece of gunk could have merely beth₁ proper parts and still be such that all of its parts themselves had proper parts (Think of a thing that has two parts of sort 1, which themselves each have two parts of sort two, which themselves each have two parts.. and so on. Such a thing would have two times two times... aleph₀ times — that is, two to the power of aleph₀, which is beth₁). But there is another conception of a way one might think gunk could be: one might think it would be possible for there to be gunk such that for any number of proper parts, the gunk would have more than that number of proper parts. One might think that, even for infinitesimally small parts, they could be divided into strictly smaller parts. Such gunk seems, to me at least, to be a metaphysical possibility — I suspect it has even been an open scientific question whether anything in the actual world is that sort of gunk. The fact that MS rules out this intuitively possible gunk is a black mark against it, though I admit that as black marks go it is a fairly minor one.

---

91 That is, no matter what cardinality of parts a piece of such Gunk is (exhaustively) divided into, there is a higher cardinality of parts such that the Gunk also exhaustively divides into those parts. Since the writing of the paper on which this chapter is based, I have coined the name “Hypergunk” for this variety of Gunk.
5. Conclusion

The realisation that Forrest/Armstrong style cardinality arguments are not successful allows us to have a simpler and more intuitive principle of recombination, and dispenses with the need to postulate mysterious necessities in order to guarantee a maximum size of worlds. As we have seen, examining the options for possibilia to form a proper class also raises new issues to do with the connection of possible worlds and classes, and may provide the motivation for various theoretical commitments that seemed previously to have no application outside pure mathematics (Should any proper classes have singletons? Should the individuals form a set?). While this essay has proceeded as a discussion primarily of Lewis’s theory, it can be more generally applied — any theory of possible worlds that would wish to endorse something like the unrestricted principle of recombination and wishes to be able to talk about possibilia from more than one world simultaneously will need a modal class theory where the possibilia form a proper class rather than a set, for example. Nearly every theory of possible worlds can, and in my opinion should, take advantage of the unrestricted principle of recombination rather than limit the size of possible worlds.92

92 Thanks to Peter Menzies, Graham Oppy, and audiences at the 1995 AAP conference and the Research School, ANU for comments and discussion on the paper which is the ancestor of this chapter. Especial thanks to Greg Restall who helped greatly with discussion and comments and pointed out a fatal flaw in an argument contained in an early draft of this paper.
Chapter 7 — A Reduction of Classes to Possible Worlds

1. How to Reduce Classes to Possible Worlds

One approach to the philosophy of mathematics is to attempt to explain or reduce mathematics or mathematical entities in terms of modal resources of various sorts. Once the restriction is lifted on the principle of recombination, as I argued it should be in the previous chapter, a proper class of possibilia become available as theoretical resources, and with these new resources a version of class and set theory can be developed which requires only the theoretical resources employed to provide for these possibilia. If such a reduction is available, it would be an important application of possible worlds, since the ontology of mathematics is a perennial problem in metaphysics. It would also incidentally provide another reason for thinking that proper-class many possibilia should be countenanced. This approach grows out of an approach to classes developed by David Lewis in his (1991), and by his co-authors John Burgess and Alan Hazen in the appendix to that book.93

For the purposes of this paper, I will employ a concretist theory of possible worlds as my example theory of possibilia, similar in most respects to that developed in Lewis (1986b), save that the restriction on the principle of recombination has been lifted. This is because this theory is in many respects the most straightforward approach to possible worlds and possibilia. However, the technique used in this paper can be applied using the resources of many other approaches to possible worlds (though not all, of course — for instance, those that take possible worlds to be sets of some sort cannot then use their worlds to reductively analyse classes on pain of circularity). For example, the various varieties of theories of possible worlds as sui generis abstracta (like Peter Van Inwagen’s so-called “Unsound Abstractionism”94 (Van Inwagen 1986 p. 201)) can be used, via their justifications of possibilist quantification over possibilia, to perform the tasks to be discussed. The main requirements of a theory of possible worlds which could provide the basis for the sort of reduction I will discuss is that the theory does not employ set- or class-theoretic devices in accounting for worlds or possibilia, on pain of circularity, and that such accounts allow for “possibilist” quantification over all of the possibilia, rather than quantifiers that are restricted to domains of single worlds at a time.

93 A version of this approach to classes is also set out in Lewis 1993.

94 For Unscientific Naive Superstitious Obscurantist Unenlightened Neanderthal Dogmatic Abstractionism (Van Inwagen 1986 p 201)
I will begin my account by outlining the approach to classes taken in *Parts of Classes* (Lewis 1991), and then go on to show how it can be adapted.

Before I begin, however, I should make a terminological note to help avoid confusion. “Atom”, as well as its chemical sense, has two distinct technical senses in set theory and in mereology. In set theory, “atom” is often used to mean “urelement”, whereas in mereology an “atom” is a partless thing (or a partless thing which itself is potentially parts of other things, which comes to the same thing in this context, but might not in a mereology which did not allow that everything is fit for mereological relationships — e.g. some mereological calculi do not allow that classes are capable of the part-whole relation, and some mereological calculi only define the part-whole relation for spaces, rather than objects in general). I will be using “atom” only in its mereological sense in this chapter and Appendix One. According to Lewis some classes are “atoms”, in this sense (the singleton classes), and many non-“atomic” objects, in the sense of “atom” I employ, are urelements of classes.

Classes, according to Lewis, are mereological fusions of the singleton classes of their members (singleton classes are classes with only one member). Singleton classes are mereologically atomic — that is, they have no proper parts. Non-empty sets, according to Lewis, are those classes that are themselves members of classes. The empty set, which according to Lewis is not a class (the only set which is not a class, in fact), is the fusion of all the individuals (i.e. the fusion of all the non-classes). This was an arbitrary stipulation for convenience, however — were we to wish to designate some atom to be the null set, or some other sort of fusion of non-classes, I doubt that Lewis’ objections would be too strenuous — though if we were to do this, we would be faced both with re-writing some of the axioms and theorems he offers, and with justifying our choice as the choice.

Two alternative accounts of the relation between singleton classes and their members are offered in *Parts of Classes*. The official line presented in the body of the book is that the relation between singletons and their members is a primitive, internal, and perhaps somewhat mysterious one, which we should nevertheless postulate because of the utility of set theory. In the appendix, however (co-written with John Burgess and Allen Hazen), *Parts of Classes* discusses a different view of the singleton relation: various strategies for reducing the singleton relation into a complicated mereological connection are discussed. Of these, my favourite is the variety which employs the

---

95 In this Lewis’ approach is very similar to the approach of (Goedel 1940), rather than approaches taken by others who admit classes, including proper classes. For instance (Bernays 1937-1954) holds that no class is identical to any set — it is just that some classes are co-extensive with some sets, while others (the proper classes) are not.
Method of Double Images\textsuperscript{96}. I will not try to explain it here, but the strategy is as follows: by means of mereology and plural quantification, one uses The Method of Double Images to be able to quantify over pairs of objects for which an ordering is specified. (That is, the Method enables us to tell which of the pair is the “1st” object and which is the “2nd”). Then\textsuperscript{97} mathematical structuralism is invoked: the idea that in our talk of the singleton relation, we are not talking unequivocally about one given relation, but are really generalising about any relation that satisfies the constraints given by the axioms governing the singleton relation. So, according to structuralism, we are talking about all the relations which have the formal character of “the relation” apparently described by the axioms of set theory. Given structuralism and given that mereology and plural quantification have served up pluralities of pairs which are effectively ordered, one then takes talk about the singleton relation to be treated as a generalisation about all the pluralities of special pairs that satisfy the conditions laid down by the axioms. It is as if we have used mereology and plural quantification to give us all the classes of ordered pairs, and then applied a structuralist interpretation to the axioms so that claims about the singleton relation are true iff the claims are true for all classes of ordered pairs which conform to the axioms. The only difference is that instead of classes of ordered pairs, we have pluralities of pairs which are “ordered” by means of the Double Images device.

This method has one great advantage, and one evident drawback. The advantage is that it explains the singleton relation using only mereology and plural quantification (resources which Lewis argues we have in any case) — and while it does not completely remove the involvement of internal relations, the two internal relations which remain — the relation of part-to-whole and the relation between pluralities and the things that make them up — are presumably less mysterious and more everyday than the internal primitive singleton relation would have been. The evident drawback is that the Appendicital account of the truth of claims about singletons is not, it seems, what mathematicians actually think they are talking about when they talk about set-membership or the singleton relation — it needs to be seen as a reform to mathematics. Lewis is reluctant to have metaphysicians tell mathematicians that they were wrong all along, and is not sure that

\textsuperscript{96} (Lewis 1991, pp 121-127), where it is credited to Burgess. I dislike the Method of Extraneous Ordering (p127-133) proposed by Hazen (and the ‘Hybrid Method’ which partially relies on it) in this context, because it is committed to there being a pre-class-theoretic relation which each atom stands in to at least one other atom. It is thus already committed to the sort of thing which the Appendix is attempting to avoid — for primitive relations holding between objects look like a special case of class-like entities, or alternatively if the relation (or dyadic predicate, or whatever) is treated as primitive and not to be cashed out in terms of entities, then it seems The Method of Extraneous ordering tries to avoid the postulation of a primitive singleton relation by postulating some other primitive relation. Requiring this relation as well as mereology and plural quantification seems distinctly worse than using only mereology and plural quantification.

\textsuperscript{97} Strictly speaking, between these two steps are further applications of the method of Double Images to provide pairs which can deal with “Gunk” — see Lewis 1991 pp 133-136 for details of this step.
how much weight his suspicion of postulated internal relations should be given, so his considered position seems to be agnostic between the view of the singleton relation given in the text and the sort provided in the appendix.

One aspect of his philosophy of mathematics which is not given quite as much emphasis in Parts of Classes, but which is quite important ontologically, is the status of the singletons themselves. We have quite a discussion of what their connection to other classes and to individuals could be. But what are the singletons themselves? According to Lewis, we know that they are atomic (Lewis, 1991, p 31). We know that there are proper-class many of them.98 If we move to structuralism, we gain a little benefit in our understanding of the singletons, but not a great deal. It becomes open to us to say that there is not anything which is a singleton once and for all, but only that one “candidate” singleton relations99 treats these objects as the singletons, another “candidate” treats those as the singletons, and so on. There is still a substantive ontological question waiting in the wings though. Whether we take some things to be singletons once-and-for-all, or not, Lewis is going to be committed to some distinctive mathematical ontology. Either way, it seems that there must be a Large number of atoms that are outside the possible worlds — since if there is a set of all possible objects, and a proper class of atoms altogether, nearly all the atoms which exist will have to be distinct100 from the totality of the possible worlds. Lewis accepts this conclusion, and the appendix says explicitly that the structuralist explanation is no help with the ontology of mathematics (Lewis, 1991, p. 142). He thinks that mathematics alone justifies the postulation of these atoms — the mysterious things that make up most of Reality.

As well as this being a pretty stiff ontological cost, these extra-worldly atoms are an embarrassment for Lewis’ theory of modality. Not being in any possible world, they are impossible objects — and it is less than ideal for a modal theory to admit the literal existence of impossible objects.101 Having both concrete possible worlds and an impossible Platonic realm is a form of extravagance Lewis should admit is a real cost.

98 By "proper-class many" I mean that there are as many as there are members of a proper class. This will obviously be the case, for there will be as many singletons that are parts of a proper class as there are members of that proper class.

99 By “candidate” singleton relations I just mean any of those pluralities of ordered pairs which satisfy the conditions, derived from the axioms of class theory, which "the singleton relation" is supposed to satisfy in Lewisian non-structuralist class theory.

100 In this paper I am using “distinct” in Lewis’ sense, so that two objects are distinct from each other iff they share no parts. For those that prefer to use “distinct” to mean “non-identical”, a word like “disjoint” or “non-overlapping” can be used instead.

101 Of course, Lewis is committed to the existence of impossible objects in any case, as the fusion of things that exist in different possible worlds itself does not exist in any possible world (Lewis, 1986b, p 211). But at least these objects resolve into parts which are all possible — which cannot be said for the “completely impossible” mathematical atoms.
Lewis does not accept that mere quantitative extravagance is anything to worry about, so the fact that mathematics commits us to many more particular things than we would have had without it is not even a cost, so the question of whether it is worth paying does not really arise. But this is a case of qualitative extravagance: mathematics commits us, according to Lewis, to a quite different kind of ontology from the sort which just commitment to concrete possible worlds would otherwise commit us. If we could have mathematics without the ontology, that would mean a theory that cost less.

It should be clear how allowing for proper-class many *possibilia* might be an advantage here. If we allow that there are proper class many things in the possible worlds, then we have enough things to provide the ontology required by Lewis’ theory of classes. If we adopt structuralism about the singleton relation, then we are nearly home free. The worlds alone provide both the ontology and the necessary mereological bedrock to ground Lewis’ philosophy of mathematics — both the singletons and the singleton relation are provided. The advantages of this approach over the picture presented in *Parts of Classes* are thus twofold: it offers Lewis the opportunity to in effect halve his ontological commitments (or halve the sort of ontological commitments he is concerned with), as well as remove an embarrassment to his theory of modality.

Of course, as I mentioned, this is not only an improvement available for Lewis. Any theory of possible worlds which takes the possible worlds to contain only individuals (however this containment, or these individuals, are to be explained) and which locates Platonistic mathematical ontology outside the worlds (and so apparently makes mathematical ontology *impossibilia*) could benefit from this reduction of mathematical ontology and ideology (provided the modal resources employed do not themselves reduce wholly or partially to mathematical ontology or ideology, of course). The use of Appendicital structuralism could also provide a better alternative to those theories of mathematical objects that locate all of the classes in every possible world, or at least claims that all of those classes whose transitive closure contains only the empty set and individuals from a given possible world are found in that world. Instead of all of this especially mathematical ontology, structuralism would require only the non-mathematical *possibilia* to serve to generate the quantity of objects needed, and so the specifically mathematical ontology, both actual and possible, could still be dispensed with.

If we do not accept structuralism, we may still be able to have the benefit offered of a reduction in the number of kinds of things, though of course the singleton relation might still be a cost. Once one accepts that size considerations do not rule out the possibility of singleton atoms being located in worlds, it is open to one to think that singletons are just another sort of thing that can be found in worlds102, and therefore on a

---

102 One problem with this sort of approach is that worlds will be composites of individuals and classes. This would be a problem if we stuck to Lewis’ approach to the fusions of classes and individuals, which
par with other *possibilia* as far as ontological cost goes. One could either have each world containing singletons for all the individuals which are parts of it (perhaps the singleton relation could be seen as the sort of relation that can bind certain things into worlds, or perhaps one could explain the singletons' presence in some other fashion), or one could imagine that one or more worlds out there are chock-full of singletons, related across worlds to individuals of various sorts by the mysterious internal singleton relation. Or perhaps singletons could be trans-world individuals (like immanent universals would be), and may be wholly present in every world in the pluriverse. There are all sorts of options open to a Lewisian here, none of which require the curious postulation of a Platonic realm outside logical space. The question “In virtue of what is it that a given singleton is part of a world?” would still require an answer, but there are many plausible replies that could be offered here. I will not discuss the non-structuralist approach further in this chapter, however, but will concentrate on the ways of employing structuralism to eliminate specifically mathematical resources.

Structuralists who take advantage of the appendix of *Parts of Classes* to remove the need to postulate the singleton relation are left with one difficulty: the problem of working out what to do with individuals. Individuals are important for two reasons — one is that many of the uses Lewis and others have for set theory involve dealing with sets of specific individuals, and the other is that, for Lewis, the fusion of all individuals forms the null set. If we were merely interested in rescuing ZF, then the individuals need not be a worry — we would only be concerned with the null set and set-theoretic constructions out of the null set, so we could easily take as candidate singleton relations any at all that satisfied the axioms (and allowed that there was at least one individual to ensure the existence of a null set — though even that restriction could be lifted if we were prepared to allow as candidates relations that assigned the role of the null set to some arbitrary atom). But otherwise, we will have a problem. What ensures that, for instance, even that objects in the actual world like you or me or my teacup are individuals? For all that we have said so far, there will be some candidate relations on which I am not an individual at all (assuming I am made from atoms) — some are ones in which not only am I a class but it is not even the case that any of my members are individuals! We would like it to come out true that certain things are individuals, but as things now stand nothing

was to deny that they ever had singletons. For it is convenient to talk of sets and classes of possible worlds (in fact, if we were to give that up, we would be back where we started), but worlds would be fusions of individuals and singletons if singletons were parts of worlds. However, there is nothing stopping Lewis allowing that some fusions of classes and individuals themselves had singletons, provided compensatory adjustments were made elsewhere. And even if we wished to keep the claim that “mixed- fusions” of individuals and singletons themselves had no singletons, perhaps allowing that a few quite alien worlds (the containers of singletons) lacked singletons would not be such a high price to pay, even if it did force us to accept that there was no class of possible worlds (though there would of course still be a fusion). Finally, of course, there is no problem with worlds entirely made up of singletons themselves having singletons, as Lewis’ claim only prohibits singletons of fusions that contain both individuals and singletons.
atomic plays the role of an individual on all the candidate “relations”. An obvious solution suggests itself — we can just stipulate which things we want to count as individuals, and then take the truths about the classes containing those individuals to be the truths that hold for all the “singleton candidates” that not only satisfy the axioms of class theory but which satisfy the stipulation as to what are the individuals.

This solution has its limits. For some purposes we might want to count all the things in possible worlds as individuals, or enough of the things in possible worlds at any rate to cause trouble. (One case might be when we want to determine the semantic value for the phrase “all the things in possible worlds”). If we do this, though, there will just not be anything left over to serve as the classes. The stipulation dividing things into individuals and classes will only work if there are proper-class many things designated as the potential classes. If we chose to keep Lewis’ two-realm picture, the atoms outside logical space postulated for the purpose of saving class theory would have been natural candidates for always being counted as classes, so this stipulation problem did not arise in such an acute form.

I believe that the problem is not insurmountable, and I am not even convinced that it is a problem. We could just hope that there are enough uninteresting nondescript atoms out there that any useful theorising we wish to do will not require taking them to be individuals. If we wish there to be a class of all possible worlds we would have to allow that some worlds containing such nondescript material will themselves have singletons, which would require some other compensatory adjustments to the system (see footnote 8). Or perhaps we should deny that there is a class of all worlds (though not deny that there is an aggregate of worlds, of course). I do not happen to like this option much (and it would rather defeat the purpose of the previous chapter), but I will not discuss it further here.

Even if the fact that, as the theory currently stands, we cannot treat all the parts of worlds as individuals is seen to be a worry, it can have a solution. What one would like would be if at least some of the atoms could do double duty — be individuals, but function as parts of classes too, ideally with all of this cashed out without requiring any distinctively mathematical ideology. The following sections of this chapter will describe one kind of method by which these atoms can do double duty.

2. How atoms can perform double duty

In this section I will show how, given proper-class many (mereologically) atomic individuals\(^{103}\) (and without any further mathematical ontology), plural quantification and

\(^{103}\) It is possible to define “proper-class many” without pre-supposing any notions of class theory (such as “proper class” — at least given Global Choice (see below). One way to state the claim that there are proper-class many atoms without presupposing class-theoretic notions is to say that the fusion of the
mereology\textsuperscript{104} can be used to provide candidates for the singleton relation, and so provide (with the invocation of structuralism) a basis for set and class theory in the manner of the appendix of Parts of Classes. The focus in the remainder of this chapter will be on solving the problem brought out at the start of this chapter: how to enable such a theory of classes to treat all of the individuals as individuals simultaneously, without running out of objects to serve as classes. In the following sections I will assume familiarity with the appendix of Parts of Classes (hereafter it will be referred to simply as “the Appendix”), as well as the main text of that book. I will also assume that the relation between singletons and classes is that of part to whole, again as in Parts of Classes — so that once the singleton relation is elucidated, an account can straightforwardly be given for classes in general. In addition, I will concentrate for the most part on providing a class theory with the assumption that there is no Gunk. This is because adding Gunk to the picture is somewhat complicated, and because the method for extending a treatment of the sort I will be discussing from a treatment of atoms alone to a treatment of atoms and Gunk has already been substantially provided in the Appendix, pp 133-136. However, I shall provide some modifications to the procedure for accommodating Gunk needed for the system being constructed at the end of this piece. I will proceed with an analogue of the Method of Double Images, as I think that the Method of Extraneous Ordering is a less satisfactory procedure for the reason outlined in footnote three.

The idea behind the approach to be taken is that atoms must serve “double duty” — the very same atoms must be capable of functioning as individuals and as classes. The challenge is to distinguish atoms to be treated as individuals (or “in their individual aspect”) and atoms to be treated as being singletons (or “in their singleton aspect”), using only the resources of mereology and plural quantification. After all, we have to produce a structure isomorphic to the traditional picture of sets, where an atom is an individual or a singleton once and for all, and it is unequivocally false to say of anything which is an individual that it has any members, and unequivocally true to say of any singleton that it has a member.

The method is surprisingly simple (or at least all that is required is a simple addition to the Method of Double Images). I will outline a proposed method of meeting the desiderata, and will then discuss what advantages and disadvantages such a system would produce.

Like Lewis, I will begin with Lewis’s “framework”: a system which employs the formal resources of plural quantification and mereology, and furthermore satisfies some constraints which can be stated without employing set-theoretic language but which

\textsuperscript{104} The systems of mereology and plural quantification to be used are those outlined in (Lewis, 1991)
provide the system with power analogous to that of systems of set or class theory. I will begin by assuming hypotheses P, U and I about the size of the universe (Lewis 1991 p 93-94). Then various “principles of the framework” need to be affirmed. Lewis simply mentions principles which he will employ as needed, without attempting to systematise them or to derive them from a comprehensive unified basis. More work could be done in providing for a basis for a framework for the job of providing for class theory, but for now I am content to follow Lewis. Principles affirmed include two Choice schemata on pp. 71-72, two Replacement schemata on pp. 91-92, a Dedekind schema on p. 88, and some principles about size affirmed on p. 90-91.

Next, designate some atom “n” (for null). Every class will contain n as a part. Let us also define some features of a relation (or quasi-relation) call the “Esingleton” (for in some ways this will be an ersatz singleton relation). I will speak to begin with as if Esingleton is a typical relation, and then provide a recasting of such talk in terms of mereology and plural quantification. Esingleton is a relation that holds between atoms and other objects — either (since for the moment we are not concerned with Gunk) atoms or fusions of atoms. Esingleton is asymmetric, and let us call the atom in the first place of the relation the “Esingleton” of the object in the second place of the relation in each case of the relation. Esingleton also obeys the following restrictions:

1. **Distinctness** — no two objects share the same Esingleton
2. **Functionality** — nothing has more than one Esingleton
3. **Domain** — n has an Esingleton, any small fusion of atoms which does not overlap n has an Esingleton, any small fusion of Esingletons and n has an Esingleton, and nothing else has an Esingleton.
4. **Null Set** — n is not an Esingleton of anything.
5. **Atomicity** — all Esingletons are atoms.

The first three of these constraints are at least roughly analogous to constraints Lewis places on the Singleton relation in his system (Lewis 1991 p. 5). Null set is distinctive in this system, but obvious enough given the intended treatment of n as the null class. Lewis’s system also has an axiom called Induction (p. 96), which derives its usefulness from classes being distinct from individuals, and so is much less useful in the system I

---

105 The adoption of these hypotheses is also the first point in the technique which relies on plural quantification: if plural quantification is not in order, or requires set theory in order to make sense, then the project is in trouble very early.

106 The names of these conditions bear some connection with the names of conditions on the singleton relation in Parts of Classes — they are often different from the ones stated there however, and the names are to suggest no more than similar functioning.

107 Atomicity and Null Set might be thought to already have been covered by what was said about the Esingleton relation in the text. Their explicit addition to the list of axioms cannot hurt, especially since it is virtuous to explicitly state axioms where possible.
am to set up. Induction has two main uses in Lewis’s system. One is to ensure that singletons are atomic (p. 96) — the counterpart of which is explicitly ensured by my Atomicity. The other thing which Induction is employed for is to ensure that the class theory he derives is well founded — that is, it has the equivalent of an axiom of Foundation or Regularity (Lewis uses the German name for such an axiom: Fundierung). This is not an essential feature of set or class theories, but it does form part of the standard axiomatisations. It would be as well to ensure that the system I will develop will also display this feature. To ensure this I must add another constraint of the Esingleton relation:

6. Foundation — whenever there are some things, either at least one of them is not an object which is the fusion of at least one Esingleton with n and which contains no non-Esingletons besides n as parts, or one of them is such that all the objects which have Esingletons among its parts are not among those things.

Note that these theses are not enough to totally characterise the Esingleton relation — for example, they leave open whether there are any small fusions of atoms containing n which lack an Esingleton. Nevertheless, they are enough of a partial characterisation to allow for a reconstruction of a notion of class which will have the familiar features (or most of them). These characterisations of classes and related terminology are only a first pass at the definitions that will eventually be adopted (structuralism has not yet been invoked, for example).

A class is any fusion that contains only n and Esingletons (n on its own is a class, but every class has n as a part). An object is a member of a class iff that object’s Esingleton is a part of that class. A class is a set iff it is small. Alternatively, we could define a set as a class which possesses a singleton, as it will turn out that given Domain and the definition of class and membership that this comes to the same thing. What to say about what individuals are is slightly tricky, as I will discuss further below. For now, assume that all and any fusion of atoms that is distinct from n is an individual.

This system provides an adequate interpretation for the axioms of class theory, and does so without any extra-possible or distinctively mathematical ontology (with the exception of n, at this stage). The versions of the “standard axioms” of set theory which Lewis recovers in pp 100-107 can also be recovered by this system, once the definition of individuals is suitably tweaked — see the appendix on proofs for these proofs.

The system of Esingletons also retains many of the intuitions which Lewis mobilises to make his account of classes plausible in the first place (these intuitions, which he expresses in the form of theses, may be found on p 4 and p 7 of Lewis 1991). His First Thesis, that one class is part of another iff the first is a subclass of the second, is true in the system outlined, as is his Priority Thesis, that no class is part of any individual. Given the assumption that all fusions distinct from n are individuals, his Fusion thesis is retained as well (though later in this chapter I will examine reasons why
the Fusion thesis ought to be restricted). His Division Thesis is also satisfied: Reality
does divide exhaustively into individuals and classes (and indeed every object does, since
all the atoms are either n or distinct from n). All of the basic intuitions he outlines in the
start of Parts of Classes are vindicated by the above system. However the above system
does not satisfy his Main Thesis (that the parts of a class are all and only their
subclasses). This might seem surprising in light of the fact that Lewis claims his Main
Thesis follows from those four basic claims. His proof that it does so is however invalid
(see Appendix Two).

The system outlined, then, provides classes which will do the jobs demanded of
them by mathematicians. It also has the advantage that, despite all atoms but n being
individuals, no individual is also a class. More modifications need to be made before the
theory will suit our other purposes, however. For one thing, the above system relies on
the Esingleton relation, whereas we wish to have a system that relies only on mereology
and plural quantification. The system as it stands also has some unusual features which
would be nice to iron out. The special status of n, for example, is somewhat of a worry.
It is a thing which is not part of any individual, and as the theory stands it seems that we
are at least committed to one piece of mathematical ontology — the null set. It would be
nice if we could dispense with distinctly mathematical ontology altogether. Another
problem is that many individuals are playing double roles which are intuitively curious,
for it seems intuitively curious that individuals are parts of, perhaps quite unrelated,
classes. It would be unusual if, for example, a fusion of atoms was a lump of rock and
also served, when fused to the null set, to be the set of all actual horses. The Method of
Double Images will be used to provide a substitute for the Esingleton relation which relies
on no more than mereology and plural quantification (plus some auxiliary assumptions),
and mathematical structuralism will come to the rescue to solve the problems of the null
set and dual roles of individual atoms.

The use of the Method of Double Images will be a relatively straightforward
device to give some “ordered” pairs such that the first one of each is an Esingleton and the
second an object. Use the Method to partition all of reality. Any R (see the Appendix, p
126) defined using the partition and which consists of b-pairs such that they obey the
conditions set out for the “Esingleton relation” will do to define a substitute for the
Esingleton relation. These Rs will be substitutes for the Esingleton relation in a manner
almost completely analogous to the manner in which defined Rs satisfy the conditions to
be singletons in the Appendix, so I will not dwell on how this is done here.

Instead of attempting the impossible task of selecting which R is the one that
corresponds with the real Esingleton relation, structuralism can provide a way to explain
talk about classes — as generalisations about any Rs which meet the requirements (as in
the Appendix, pp 140-141). In fact, structuralism should be taken further. There are an
infinite number of choices which could be made about which atom is to be treated as the
null set — which atom is to receive the honour of the name “n”. In the absence of any
distinctively mathematical ontology, selecting which atom is really n is as hopeless a task
as determining which R is the real R in the absence of any distinctly mathematical
relations. The candidates to jointly-satisfy the conditions we have put on the Esingleton
“relation” and the atom to be treated as n will be pluralities where all but one of them are
b-pairs of the sort discussed in the previous paragraph and one of them is an atom — the
atom which is to be treated as n relative to those b-pairs. Each plurality (let us call them
Ps) which has as one of them an atom and the rest of them as b-pairs of the sort described
will satisfy the formal conditions to have the atom taken to be n and the plurality of b-
pairs to be taken to be the R which will do the work of the Esingleton relation. Relative
to a given P, then, it will be possible to state which atoms are members of which classes,
and which fusions are classes and which are individuals, and so on. A statement in set
theory (or class theory) will, according to structuralism, be true iff it is something which
is true for all P. The null set possesses a singleton, for example, because no matter
which P we select, the atom designated to be n will have its image be the “second”
member of a b-pair (and only one b-pair, from Functionality), and so the object which
has its first image as a member of the b-pair will count as the Esingleton of the atom n
relative to that P. There will be a fusion of that first member and the “n-atom”, which
will count as the singleton of n relative to taking the given P to give us the Esingleton
“relation” and the null atom.

Let us suppose that structuralism also holds that everything in class theory that is
not true according to all assignments of P is false. Other approaches can be taken to those
claims that are true of some-but-not-all Ps — one could take them to be truthvalueless, or
of indeterminate truth value, or true-to-a-degree less than one, and so on (the range of
options are in some respects similar to those offered to someone tempted by
supervaluational approaches to vagueness). Taking them all to be false is probably the
most natural way to treat them, if we take statements to be treated structurally to be
implicit generalisations about all structures of a certain sort — for if a claim about all
things of a certain sort is true of some but not others, we are usually happy to assign
“false” to that claim. This will produce some oddities — while it will be true, for
example, that some atom is the null set, it will be false, for any atom that we select, that
this particular atom is the null set. These counter-intuitive results are part of the price for
structuralism, but it is not a great price once we are aware of the implicitly general nature
of the claims, and especially if structuralism about sets and classes is to be seen as a
reform to mathematics rather than merely extending or clarifying it.

One complication might be worthwhile in order to preserve more of our intuitions
about what should be called individuals. If something had to be treated as an individual
by all Ps before it was true that it was an individual, then nothing constituted from atoms
could be an individual, since for each atom there will be an assignment which assigns it to
be the null set, and assigns the other atoms in the object as Esingletons. A revision which
held that nothing was an individual, but also that for almost each thing, that thing was
neither a class nor an individual, would be a paradoxical-sounding revision that differed
quite sharply from our pre-theoretic intuitions, and would also make class theory
unsuitable for many applications where it is assumed that certain objects are individuals.
So a better way of assigning the title "individual" should be developed if possible. And
such a method is easy enough.

Instead of saying that an object is an individual iff all Ps assign it the status of an
individual, it would be better to say that an object is an individual iff it is not a class (i.e.
there is some P which does not assign it the status of a class). Then it would turn out that
everything would be an individual, while it still being the case that there are enough
classes to go around. Another problem concerning individuals is that each non-atomic
individual will have the null set as a part but not have all of its other atoms being
Esingletons according to some P or other, and so will not count as being members of
classes according to that P, and so not at all. This should also be avoided. One of the
most convenient method of avoiding it when dealing with a specific bunch of objects one
wishes to treat as individuals is to generalise only about those Ps according to which all
the objects in question are distinct from the null atom. Of course this will not work when
we wish to treat everything (or even every small thing) as being both individuals and
members of classes — or even when we wish to treat of things whose aggregates are all
of (atomic) Reality.

This is therefore only a partial solution to the problem of trying to provide for all
of class theory without embracing specifically mathematical ontology, especially since the
problem of dealing with Gunk has not yet been tackled. It is an improvement over the
solution suggested initially at the end of section 1 of this chapter, in that the solution
offered there required that a proper class of atoms be distinct from the objects being
considered, while the solution sketched above requires only one atom be distinct from the
objects to be treated as individuals (and can even overlap some of them, though of course
there must be a proper class of things that the atom does not overlap). Even this solution
can (and will) be improved, but before I move on to consider improvements, it will be
worthwhile to consider some of the features of the current approach, and some features of
approaches which seek to have individuals do the work of classes generally.

3. A Surprising Feature of the Account

One surprising feature of this system of employing individual atoms to serve also as
Esingletons is that not every fusion of individuals can itself be an urelement of a class.
Here is a proof: suppose that there were proper-class many atomic individuals, and any
fusion which divided exhaustively into some or all of those atoms was a member of a
class. By the principle of unrestricted composition, there would be a fusion of those
atoms corresponding to each sub-class of the class of atomic individuals. For every P, if all of the fusions were members of classes, they would all have Esingletons — so the sub-classes of the proper class of individual atoms could be put in one-one correspondence with the Esingletons of individuals composed entirely of those atoms. Since the Esingletons are the members of the proper-class of atomic individuals, the sub-classes of that class must be able to be put into one-one correspondence with the members of a sub-class of that class. This can be shown to be impossible, using a variation of the argument for Cantor’s theorem:108

For any given P:

Let the Fs be the pairs, each of which contains a sub-class of the proper-class of individual atoms (call the plurality of these sub-classes the Xs) and an Esingleton of an individual which exhaustively divides into the members of one of the Xs (call these Esingletons the Es), such that each F has as one of its “members” the Esingleton of the individual formed by the fusion of the members of an X which is also one of that F. Consider the Esingletons which are those of the Es that are not one of any F such that they are a member of the X which is one of that F — Call these Esingletons the Ks. Consider the fusion of the Esingletons which are members of K. Is the Esingleton of that fusion (call it x) one of the Ks? Either way lies contradiction:

Assume that x is one of the Ks. Then it will be identical to none of the members of the X which it is paired with in an F (as that is what it is to be one of the Ks). The X which it is paired with, however, will be the class of all and only the atoms which are the Ks (by the definition of the Fs). So, in particular, one of them will be x (ex hypothesi). So x will not be identical to itself.

Assume that x is not one of the Ks. Then it will be identical to one of the members of the X which it is paired with. The X which it is paired with will be the class of all and only the atoms which are the Ks. But x is not one of the Ks, by hypothesis, and so will not be identical to itself.

Since this reasoning holds for every P, it follows that, given a proper-class of individuals that are members of classes, not all the fusions of those individuals can be members of classes, on pain of contradiction.

For this proof to work, I need to employ a sense of “pair” according to which proper-classes can be “members” of pairs. The standard set-theoretic method of taking pairs to be sets of various sorts will not therefore be adequate, since proper-classes are never members of sets. The pairs cannot be pluralities of two either. Since the proof plurally quantifies over the pairs, the pairs cannot themselves be pluralities, since

108 Thanks to Greg Restall for pointing out that this sort of proof can be employed using plural quantification even when functions are unavailable.
standard plural quantification (see for example Boolos 1984) is characterised by the lack of ability to plurally quantify over the pluralities themselves, but rather allows only plural quantification over the objects which make up those pluralities. It would be possible to construct systems that allowed for higher-order plural quantification over pluralities (as is discussed in Hazen 1996b), and the proof could be construed as plural quantification over two-membered pluralities in such systems, but I shall not avail myself of such a device. Apart from my suspicion of such a device, higher-order plural quantification is not something which is part of the framework of Lewis's in which I am conducting the present investigation. Thus, I would prefer a device for producing “pairs” which did not rely on resources not available in the framework. Fortunately, there are resources available in the framework to provide pairs of any two objects, even proper classes: the Burgess method of producing pairs (b-pairs) and the Hazen method of producing pairs (h-pairs). Since I have been employing the Burgess method, let me do so here — and take the “pairs” in the above proof to be the b-pairs of sub-classes and Esingletons of the appropriate sorts, where the b-pairs are the b-pairs according to some arbitrary suitable \(x, y, z, X, Y\) and \(Z\). With a suitable device for producing the “pairs” of the above proof, it goes through, and so establishes the conclusion that if there are proper-class many atomic individuals which belong to classes, then not every fusion of those individuals is itself an urelement of a class.

In light of this result, it may be worth refining our conception of “individuals”. Our current definition is that any non-class is an individual. Given this definition, it follows that in this system not all individuals are members of classes. This conflicts with more normal usage of the term “individual” (for example in Fraenkel, Bar-Hillel and Levy, 1973, p 23) in set theory, which is that an individual is any member of a class that is not itself a class. Let us adopt this definition instead (keeping in mind that the definition is to be applied after generalisation rather than before). Thus, instead of saying that some individuals are not members of classes, we will say that some fusions of individuals are not themselves individuals. This might strike some as worrying, or even absurd. It is not as bad as it first appears, however.

Most of the fusions of proper-class many individual atoms will need to be denied the status of individuals themselves, and once singletons are denied to these entities (or most of these entities), the other fusions of individual atoms can all be admitted to be individuals without causing any trouble. This seems to be the obvious approach to take in the light of this result, and it ought not to be too surprising that the counter-intuitive results of this theory happen in connection to things having proper-class many parts: the paradoxes of set (and class) theory have been known (or suspected) to stem from some postulated things being “too big” since the time of Cantor.\(^{109}\) The analogy is even closer

\(^{109}\) Hallett (1984) is a good discussion of the connection between “size” and the paradoxes.
if the relation of sub-class to class is taken to be mereological: for just as a proper class has no singleton in virtue of having proper-class many singletons as parts, so a Large non-class lacks a singleton in virtue of having proper-class many non-classes as parts. The denial of singletons to all aggregates of individual atoms is of course consistent with unrestricted mereological composition: aggregates of all and any individuals do still exist according to this theory, it is just that some are not members of classes. Furthermore, if it is stipulated that, out of the fusions which exhaustively divide into individuals, only the fusions of proper-class many atoms may fail to have singletons, the “Restricted Fusion Thesis” that the fusion of any set of individuals is itself an individual follows from the current system. However, Lewis’ full Fusion Thesis (that any fusion of individuals is itself an individual (1991 p7)) will need to be denied.

If one wishes to retain the full Fusion thesis, one will need to abandon the project of employing proper-class many individuals to serve as the ontology for class theory. But, given a plausible class-theoretic assumptions, this is not because the atoms are being pressed into serving double duty: the denial of the full Fusion Thesis is a direct consequence of the postulation of a proper class of individual atoms. This means that, given that one wishes to postulate a proper class of individuals (as for instance I argue in the previous chapter that Lewisian modal realists should), postulating the usual class-theoretic ontology on top will not help, so it has to be seen as a cost of postulating the individuals in the first place, not some further problem one becomes embroiled in when one tries to reduce classes away. The proof that it follows from their being proper-class many individuals is not as straightforward when we countenance the prima facie option that there might be more classes than even proper-class many individual atoms (an option which was clearly ruled out when the classes were one-one correlated in the obvious fashion with E-singletons each of which was also an individual atom).

The above proof showed that the fusions of individual atoms could not be put in a one-one correspondence with those atoms. The additional piece of information to be used in the proof that the singletons cannot be put in one-one correspondence with the fusions of individuals atoms when those individual atoms form a proper class will consist of a proof that all proper classes are of the same size: they are all able to be put in a one-one correspondence with each other. Once this is established, it will follow that the singletons can be put into one-one correspondence with the individual atoms (given that the singletons obviously form a proper class), and so cannot be put in one-one correspondence with all of the fusions of those individual atoms: thus, (given Functionality) some of those fusions will lack singletons.

All that need be shown now is that all proper classes are of the same size. Lewis is already committed to this (Lewis 1991 p 98). This claim is embraced by nearly all theorists who accept a theory of classes: von Neumann made a claim which is equivalent
to this one of the axioms of his class theory (axiom IV2 of von Neumann 1925)\textsuperscript{110}. The claim that all proper classes are of the same size is equivalent to the axiom of global choice\textsuperscript{111} (as it is called in Fraenkel, Bar-Hillel & Levy 1973, p 133), which they argue I think persuasively is a quite reasonable assumption when dealing with classes, and which is taken by Rubin and Rubin 1973 to be the natural axiom of choice to accept in class theory (were one to embrace any axiom of choice). Global choice is by no means merely an artifact of using the same ontology to do the work of individuals and classes — it is a fundamental principle of class theory. While of course there is no inconsistency in denying it, and indeed denying global choice is even consistent with choice (See Fengler 1976 p 278), the fact that the full-blown Fusion Thesis when there is a proper class of individuals is equivalent to the denial of Global Choice makes the Fusion Thesis unattractive in its own right in such cases.

4. Utilising The Surprising Feature

Denial of the unrestricted Fusion thesis is unusual and perhaps somewhat regrettable, but since it only applies to fusions of proper-class many atoms, which is not the first time that size has led to intuitively odd results in class theory, and since it is not an unintuitive consequence that holds solely because of the double duty objects are serving in the theory but would arise in any theory postulating proper-class many individual atoms which had the axiom of Global Choice, this result does not I think damn this reductive project. In fact, once it is recognized that fusions of proper-class many individual atoms may not themselves be individuals, this fact can be utilised to iron out the bug mentioned on p 159. The bug, remember, was that we were unable to talk about too many things as individuals at once, for it had to be the case that there was at least one P such that all of the individuals being discussed were distinct from the atom designated as n by that P. This bug can be corrected by using one of the large fusions to fill the same sort of roll that the atom n played — or, to be more precise, having, for each equivalent of a P, one of the large fusions of individuals play the role which is played in the Ps by atoms playing the “n” role. This will make it possible to talk of the class of all individuals, as we shall see, in a way which the previous system outlined could not. To outline how this is to work, I will follow the procedure used earlier, of first talking as if there is only one plurality of

\textsuperscript{110} As cited in Fraenkel, Bar-Hillel and Levy 1973 p 137, where they provide a recasting of the idea as axiom (\textasteriskcentered). (\textasteriskcentered) directly implies that if a class is not a set then there is a function which maps it onto the class of all sets: therefore, every proper class can be mapped onto the class of all sets, so any proper class can be mapped one-one to any other.

\textsuperscript{111} see pp 84-85 of Rubin and Rubin, 1963, for a proof that all proper classes are equivalent in size given the Well Ordering Theorem for classes (i.e. that their WE SS or their P 1S is equivalent to their WE 1S), and pp 86-89 for a proof that the Well Ordering Theorem is equivalent to Global Choice (their
things which satisfies the criteria for being treated as the basis for talk about classes, and then generalising over all structures that meet the postulated conditions to produce the structuralist version of the theory.

As before, the background for this system will be the Framework, with the various Schema, affirmed principles, and megethological postulates about size which I mentioned on p 164 of this chapter. Begin by selecting a fusion of proper-class many atoms such that it is completely distinct from at least one other fusion of proper-class many atoms. Let this big fusion be the "Null thing" — symbolised N. As before, an Esingleton relation can be postulated. Distinctness, Functionality and Atomicity can remain unchanged, while Domain and Foundation only need to be altered slightly, as follows:

**Domain** — N has an Esingleton, any small fusion of atoms has an Esingleton, any fusion of a small fusion of Esingletons and N has an Esingleton and nothing else has a singleton.

**Foundation** — whenever there are some things, either at least one of them is not an object which is the fusion of at least one Esingleton with N and which contains no non-Esingletons besides parts of N as parts, or one of them is such that all the objects which have Esingletons among its parts are not among those things.

There need be no axiom analogous to the previous axiom of Null Set, since N will not be atomic and so cannot be an Esingleton by Atomicity. In addition, add this principle:

**Esingleton Distinctness**: every Esingleton is distinct from N. (That is, none of the atoms (or, indeed, anything) which are parts of N are Esingletons).

Since there are at least proper-class many atoms distinct from N (from the definition of N), there are sure to be enough Esingletons to go around even with this condition.

Definitions much like those on pp 164-165 can now be employed. A *class* will be any fusion that contains only N and Esingletons (N will be the null class, in the same way that n was). An object is a *member* of a class iff that object's Esingleton is a part of that class. A *set* can be defined as any class whose mereological difference from N is small. (The previous definition, that a set is a small class, is no longer adequate because all classes will be large in virtue of having N as a part). Let us say that an *individual* is any small object.

The "standard axioms" are again recovered (see the appendix on proofs for the required proofs). Again, Lewis's Main Thesis is falsified, but the First Thesis, Priority, Division and the restricted Fusion thesis (see above, p 158) are all satisfied by this system.
To dispense with the "Esingleton relation", and to iron out problems with this system like many individuals being part of the Null Class, we can now modify this theory using the Method of Double Images in the same way that the initial account of classes presented in this piece was modified. Any $R$ defined using the Method of Double Images which consists of $b$-pairs such that they obey the conditions set out for the "Esingleton relation" will do to provide a structure with $b$-pairs that can obey the conditions set out for the "Esingleton relation". Before the structure is entirely adequate, of course, some large fusion needs to be assigned the status of $N$ (and in fact this must be done before it is determinate whether a plurality of $b$-pairs satisfies the Esingleton conditions, for they must not assign to anything an Esingleton which is a part of $N$). So, as before, we will use pluralities (call them the $P$s again) such that one of each of them is a large fusion satisfying the conditions on $N$ and the rest of each of them are diatoms generated using the Method of Double Images and obeying the conditions set out on Esingletons (treating the relevant large object as $N$). Then claims about classes can be taken to be generalisations about all the $P$s satisfying those conditions: a claim about classes is true iff it is satisfied by all $P$s, and false otherwise (though again, as on p 63 different approaches can be taken to the "true of some-but-not-all" cases). Individuals can then be defined as any small object which not all $P$s assign the status of a class — that is, any small object. If one wishes, one can allow that some large objects are individuals too, in which case extra conditions should be added to the Esingleton conditions to ensure that such objects are always assigned Esingletons and are never classes.

We now have a system that is nearly ideal. Every small object is an individual, and yet every small object is such that it has a singleton. There is a class of all individuals — indeed, there is a class (the union of all classes) which has every member of a class as a member of it. Furthermore, all of this is modelled without any distinctively mathematical ontology or ideology. There are of course some drawbacks. The structuralism yields some curious consequences: for example, it is true that there is some atom which when fused to the null set yields a class, but for each atom it is false that that atom, when fused to the null set, yields a class; and the null set is a fusion of individuals, but for any given individuals it is false that exactly those individuals when fused yield the null class; and so on. But this is a feature of structuralism in general, and is a consequence that is tolerable once it is understood what is really going on. There are some lurking entities which look strange — there are things which are fusions of individuals but which are not themselves individuals, and for most $P$s there are fusions of classes and individuals which are themselves neither classes nor individuals (but this is falsified by the $P$s in which Reality exhaustively divides into $N$ (or rather the $N$ candidate) and Esingletons (or rather the Esingleton candidates). But such objects only appear amongst the Large — and the Large are in an area where our intuitions must be modified away from our naive preconceptions in any case.
I have so far been setting to one side the question of what to say in the case in which Reality is not completely atomic, but also contains Gunk, but extending the systems offered requires little that is novel. The system set out on p. 133-136 of the Appendix will be completely satisfactory when the fusion of Gunk is small. The only alteration that is required to the system is that the pairs which are to appear in the Ps should be new pairs (in the sense stated on p. 135 of the Appendix), and each of the Ps should have three marker atoms. When there are proper-class many pieces of Gunk, the procedure becomes a little more involved, for the proof offered for the “Not-too-much-Gunk Hypothesis” on p. 134 of the Appendix will then not work. In cases where there are as many pieces of Gunk as there are atoms, there will still not be a problem — The Not-too-much-Gunk hypothesis will still be satisfactory.

However, a problem might arise similar to the problem of fusions of proper-class many atoms if there is too much Gunk. If there are proper-class many distinct pieces of Gunk, then there will be more fusions of Gunk than there are members of a proper class. Part of the solution is the same as the one outlined before — it will turn out that not all large fusions of Gunk will be members of classes — i.e. not all will have singletons. This will require modifying Domain so that instead of talking about “any small fusion of atoms” it talks about “any small fusion”. As before, we do not have to deny singletons to every Large piece of Gunk, so Domain could allow some certain specified Large fusions of Gunk have singletons as well. For instance, if it is a possibility that there is the sort of Gunk described on p. 153 of this thesis — Gunk such that for any size the Gunk divides exhaustively into parts of that size — then one may wish to declare that maximally connected pieces of such Gunk have singletons.

But this will not solve the problem entirely. For the method for assigning codes to pieces of Gunk and objects which are part Gunk and part atomic requires that there be as many atoms as there are pieces of Gunk (see the Appendix, p. 135). We need a modified method of assigning codes. Fortunately one will be possible, since we do not need codes for every piece of Gunk — only those pieces which we will desire to have singletons will need to be coded so that their codes can appear in the new pairs which will be found among the Ps. To define the Gs which will be used to define the required codes (see p. 135) we need only replace the Not-too-much-Gunk Hypothesis with something like the

**Not-too-much-relevant-Gunk Hypothesis:** There are some things G such that each one of G is the fusion of a small atomless thing and exactly one atom; every small atomless thing is the maximal atomless part of exactly one of G; and no two of G have an atom as a common part.

Of course, the Not-too-much-relevant-Gunk Hypothesis can and should be modified if it is desired that some large pieces of Gunk are to have singletons — then instead of “small atomless thing” a phrase of the form “small atomless thing or large
atomless thing such that... (fill in desired condition here)'' should be substituted. Care must be taken that not too many Large pieces of gunk are assigned codes, but this is a pitfall which is easy enough to avoid. Once the codes are assigned in this fashion, the procedure can be carried out as before. The Not-too-much-relevant-Gunk Hypothesis is still a mild restraint on the amount of Gunk there is — there cannot be more small fusions of Gunk than there are atoms, after all, but like the case of fusions of atoms this is not something peculiar to the reductive project being discussed in this piece — if all of the small fusions of Gunk are to have singletons, and if singletons are atomic, then obviously there cannot be more small fusions of Gunk than atoms regardless of whether the singletons are taken to be part of a Platonic ontology or reduced to a non-mathematical ontology of individual atoms.

Perhaps even more than this can be accommodated in a system with the broad outlines of a framework of the sort dealt with in this chapter or in Lewis 1991. Hazen 1996b points out that there is a formal result, Stone’s theorem (see Hazen 1996b pp 5-6), which shows that any non-atomic mereology can be isomorphically mapped into an atomic mereology. As Hazen goes on to argue, such mappings could be constructed, pieces of gunk could play the roles which the Method of Double Images and the axioms of Esingleton reserve for atoms, and so Gunk could be fitted into a framework for class theory without needing to be associated with atoms in the manner of the Appendix. I say only perhaps: for the use of Stone’s lemma typically involves the employment of powerful set-theoretic devices such as ultrafilters (Hazen 1996b p 6), and so Stone’s theorem as it stands cannot be relied upon without circularity. Hazen suggests that there are extensions to the framework which will allow Stone’s theorem to be employed without appeal to set-theoretic entities: Hazen’s specific suggestion is higher-order plural quantification. I doubt if I have much useful to say either for or against such higher-order plural quantification — however, it does not strike me as having the harmlessness of standard plural quantification, and that I am suspicious that it is smuggling theoretical resources which are the equivalent of set theory back into the pre-set-theoretic framework. My expression of opinion here should not be taken as an argument, of course.

Alternatively, we could simply add as an assertion that the requisite pairings for the job exist (Hazen outlines what pairings are needed for a non-atomistic framework in Hazen 1996b pp 3-5). They seem fairly harmless, and their existence may even follow from suitably strong forms of Plural Comprehension, without a detour through a proof of an equivalent of Stone’s theorem. More work needs to be done on the question of how well the existence of the necessary pairings is justified, and how Esingleton are to be characterised in a non-atomic system: and I will not attempt a full discussion of this here. Apart from anything else, there is probably no need: the modal universe is almost certain to have enough atoms for the versions of the framework relying on atoms to work (we
would be in trouble if mereological atoms could be shown to be metaphysically impossible, but there is no such proof in the offing. Nevertheless, extensions to non-atomic frameworks provides an interesting technical problem if nothing else, and may well be worth exploring if only to give the framework a greater deal of generality.

Let me note as an aside that as well as extending the system in various ways to fine-tune its treatment of Gunk, there are other extensions which are possible and may perhaps even be desirable. The fact that not every Large fusion of individuals can itself be an individual does not by itself preclude *all* of the Large individuals from membership of classes, and with suitable modifications of Domain certain privileged Large individuals may be accorded the honour of a singleton. One case might be if one had an ontology which included Large cosmoi or possible universes — for some purposes it might be convenient if every possibly-maximal *possibilium* was a member of a class, perhaps for similar reasons to the desirability of every possible world belonging to a class. A similarly relaxed attitude to allowing singletons to some Large classes (the classes normally thought to be the proper classes) could also be taken — some reasons why it might be desirable that some proper-classes themselves be members of classes can be found in section 3 of chapter 6. Again, not every proper class can be a member of a class on pain of Cantorian paradox, and as usual certain proper classes should be forbidden singletons so as to prevent paradoxes analogous to the Russell paradox or the Burali-Forti paradox. In the first case, one had better forbid a singleton to the class of all non-self-membered classes which themselves are members of some class, and in the second one had better take care that the class of ordinals cannot belong to a class of higher ordinality. Again, Domain would have to be modified, and whether such tinkering is worth the bother will often be contentious philosophically. Such options for expansion will not be examined further in this chapter — I leave them for those who have some specific use for them.

With the accommodation of Gunk, the reductive system outlined reaches its final form (or at least the final form it will assume in this thesis). It does as well as any Platonic rival, save from the oddities that it inherits from being structuralist and from postulating a proper class of individual atoms. However, it requires no specifically mathematical ideology, and no distinctively mathematical ontology either, especially for a theory already committed to being able to make useful sense of possibilist quantification without doing so in terms of mathematical ideology and ontology. As such, it is a valuable contribution to the project of using modal ontology and ideology to provide the foundations of class theory and therefore, perhaps, of mathematics.
Recapitulation

As I mentioned in the introduction, the chapters of this thesis are connected loosely. Let us nevertheless see what has been achieved, and how the results argued for in this thesis might fit themselves into a larger project.

In the first chapter, two senses of the term “possible world” were distinguished: the sense in which possible worlds are thought to be cosmoi or universes: much the same sort of thing as the cosmos or universe of which we are a part; and the sense of possible world as objects which play certain theoretical roles and are correlated to modality in the relevant systematic way. At least a rough outline was given for the conditions which a candidate group of objects (or more correctly, a (group of objects, relation) pair) would need to satisfy to be entitled to the title of possible worlds, in the second sense I distinguished. The conditions (the correspondence condition, the naturalness condition, and the perspicuous connection with modality condition) were probably not terribly surprising, and to the extent that my discussion of the concept of possible worlds was meant to be an explication of the notions in use, it would have been a problem if the explication had been too surprising. Nevertheless, explicit statement of the desiderata will be helpful in evaluating the host of rival accounts of possible worlds which are being offered (and the far greater hosts which could be potentially offered). The taxonomy, while being broad-brush and possessing no great degree of originality, might also serve to make more explicit the range of options available to theorists of possible worlds (especially those who are temperamentally realists about the truth-conditions of claims about possible worlds, in the sense of “realist” I explained).

The second chapter attempts to address the issue of how we are to assess theoretical primitives, as one of the first steps towards providing standards by which to compare rival theories. Both the more general issue of how we are to assess the theoretical primitives of theories in general, and the specific issue of how it is to be done in theories of possible worlds and modality, are issues which could do with more examination than they seem to receive in the literature. Chapter Two at least makes a start on what seems to me to be a pressing philosophical need to do more work in this area: hopefully it makes enough progress so that the harder issue of what the theoretical virtues and vices are which relate to the number and nature of a theory’s primitives can also receive more attention. (Attention which I hope to provide in future work).

In the third chapter, I deal with one of the important arguments for claiming that possible worlds should have primacy relative to modality. In providing modal-operator translations of the Hazen cases, and arguing that the second order translations employing modal operators are the natural approach to a class of ordinary language claims in which the “challenging” Hazen cases (cases five and six), I provide an alternative to dealing with
these cases by means of possible worlds. The fact that this can be done does not, perhaps, prove anything conclusive about the relative primacy of modal operators and possible worlds (perhaps because the Hazen cases did not putatively show anything conclusively about this issue either), but insofar as a “modal-operator first” theorist may have wished to argue that all of the modal truths could be captured in principle without use of the modal resources of talk of possible worlds, Chapter Three offers that ability with respect to the Hazen cases at least. Of course the issue raised by chapter three, especially the first half, is the status of neutral quantification for the modality-first theorist. Allowing for such devices while sailing between the Scylla of Meinongian metaphysics and the Charybdis of possible-worlds-first theory is another project for the future (and another, unsurprisingly, which I look forward to addressing).

Chapter Four raises four problems for that variety of modal fictionalism which seeks a reductive explanation of modality in terms of a fiction we employ about possible worlds. The objections, that strong modal fictionalism would make modality too artificial; that the strong modal fictionalist has difficulties providing a modal fiction which represents all that she needs it to without an appeal to modality; the difficulty of accounting for propositions; and the worry that any constraints used to ensure that the fiction selected is the correct one will themselves provide the basis for a more direct account of modality; together make the strong modal fictionalist project very unattractive, and quite possibly untenable. Even the invocation of *sui generis* propositions, with the costs that this brings, does not entirely enable the strong modal fictionalist to evade these problems (or analogues of these problems), and besides makes the strong modal fictionalist’s position close enough to that of linguistic abstractionists that it becomes hard to see why the strong modal fictionalist’s position is to be preferred. If I am right, then this version of a possible-worlds-first theory is not acceptable — which goes a little way (though only a little way) towards showing that theories which take modality to be prior to possible worlds are to be preferred to the alternative. Timid modal fictionalism, the version of modal fictionalism which does not take the account of the acceptability of possible worlds talk to provide an account of modality, does not face the problems facing the strong modal fictionalist to the same extent, and this fact should also be encouraging to someone like myself concerned to avoid accounting for modality in terms of possible worlds.

In dealing with the problem of talking meaningfully about alien universals in Chapter Five, in particular as this problem arises for modal fictionalists and linguistic abstractionists, I defend two important actualist accounts of possible worlds against one of the serious objections which face them. My solution, in terms of a massively inter-defined world-book, is a novel approach to such theories, and so another effect of this chapter, if it is correct, is to modify the standard picture of stand-alone world stories. It
is thus potentially an important development in two of the most popular styles of accounting for the truth or acceptability of claims about possible worlds.

Chapter Six is concerned with arguing for the unrestricted principle of recombination: that for any objects in any worlds, there is a world which contains any number of duplicates of all of those objects. In showing that the standard Forrest/Armstrong argument which motivates restricting this principle of recombination fails, and providing a new argument which, I believe, is the best that can be done to come up with a similar conclusion to the Forrest/Armstrong argument, I show that there is no compulsion to restrict the principle of recombination. However, if it is not restricted, I argue, this will mean that there will be a proper class of *possibilia*. I argue that this is not undesirable, and in doing so I provide a potential modification or extension for many of the standard theories of possible worlds.

This modification to an account of possible worlds is put to good use in Chapter Seven, where I show that, provided one has a proper class of *possibilia*, one can produce a reduction of class theory which dispenses both with any distinctively mathematical ideology and also any distinctively mathematical ontology. It thus provides a new answer to the question of the nature of the truths of mathematics, an important issue which has exercised most of the major figures of analytic philosophy. Many will not find it acceptable, not least because of the controversial nature of the resources it employs, but I hope that it may form some small contribution to the investigation of mathematics. It is only one of many accounts that can be given of the connection between modality and mathematics, and I hope to do further work in examining the options when it comes to the connections between modality and mathematics.

This thesis resembles a collection of jigsaw puzzle pieces more than a detailed canvas, and it is clear that, even if every argument in it is sound and every position taken correct, much more research in the philosophy of modality and possible worlds is warranted. Fortunately enough, I suppose — it is an interesting and important area in which to work, and I would not want all the problems solved too soon. However, this thesis addresses enough important issues and provides enough important results so that it is a worthwhile contribution to the greater task.
Appendix One: Proofs for Chapter 7

Section 1: Proving that my systems satisfy Lewis’s so-called “Standard Axioms”.

These proofs are not terribly original: they are closely modelled on the proofs Lewis gives for the recovery of the “standard axioms” of class theory, including often the same wording. I include them only to show that the results I claim for my system can in fact be derived. The only radical deviation is that Fundierung is derived from my Foundation instead of an analogue of Lewis’s Induction, since the analogue of Induction in my systems does not deliver the desired result, and would lack intuitive support.

Another difference between my proofs and Lewis’s worthy of mention is that I specifically invoke the existence of an infinity of atoms, whereas in the corresponding points in Lewis’s proof he leaves that implicit. There is a proof from Lewis’s Domain, Functionality and Distinctness, plus a result which he derives that singletons are atomic, plus his version of the Dedekind Schema (an account of infinity affirmed as a principle of the framework on p. 88-89, to the conclusion that something is infinite). A similar result can be derived from my Domain, Functionality, Distinctness and Atomicity, plus the Dedekind Schema. I will show this result (using virtually the same argument as Lewis), and then refer to it as IR (for “infinity result”) in the further proofs.

In producing these proofs, the materials I will employ are some of the materials which Lewis invokes in setting up his “framework”: the rules of mereology and plural quantification, which Lewis provides partial axiomatisations for on pp 73-74 and p 63 respectively; the three “hypotheses of size” P, U and I (pp 93-94); the Replacement Schemata (Lewis 1991 pp 91-92); the Choice schemata (pp 71-72); the various affirmed “principles of the framework” given by Lewis: “if there is something infinite that consists entirely of atoms, then (1) any finite thing is small, and (2) any fusion of a small thing with a finite thing is small”, from p 90-91, and “something is small iff its atoms are few”, from p 91; and finally the definitions he provides in chapter 3 for the distinctions of size on pp. 88 and 89. Given these, and given the axioms I provide for the various Esingleton relations I discuss in chapter 7, and given my definitions of class-theoretic terms, the standard axioms follow.

I will quote the various schemata which form part of the framework when they are needed in the proofs below, and the other principles can be found scattered through Lewis’s book. Since there are two proposed sets of axioms for the Esingleton relation (not counting structuralist variants) in chapter seven, there will correspondingly be two sets of proofs of the “standard axioms” which I shall provide in this appendix.
Proofs for my first proposed system

AXIOMS AND DEFINITIONS

For the first proposed system in chapter 7 about which I make the claim that the axioms of class theory which Lewis calls the “standard axioms” can be recovered:

The axioms of Esingleton:
1. Distinctness — no two objects share the same Esingleton
2. Functionality — nothing has more than one Esingleton
3. Domain — n has an Esingleton, any small fusion of atoms which does not overlap n has an Esingleton, any small fusion of Esingletons and n has an Esingleton and nothing else has a singleton.
4. Null Set — n is not an Esingleton of anything.
5. Atomicity — all Esingletons are atoms.
6. Foundation — whenever there are some things, either at least one of them is not an object which is the fusion of at least one Esingleton with n and which contains no non-Esingletons besides n as parts, or one of them is such that all the objects which have Esingletons among its parts are not among those things.

The definitions of class-theoretic terms:
A class is any fusion that contains only n and Esingletons (n on its own is a class, but every class has n as a part). An object is a member of a class iff that object’s Esingleton is a part of that class and no non-class has any members. A class is a set iff it is small. What to say about what an individual is slightly tricky, as I will discuss further below. For now, assume that all and any small fusions of atoms that is distinct from n are individuals.

PROOFS

IR: there is an infinite fusion of atoms.

This proof will employ the Dedekind Schema affirmed as a principle of the framework on p. 88-89 of Lewis, which is as follows:

If x is a proper part of y, and if each atom of y ... exactly one atom of x, and if each atom of x is such that exactly one atom of y ... it, then y is infinite.
Proof. By Domain, there are Esingletons, for instance, the Esingleton of \( n \), and all diatoms consisting of \( n \) and an Esingleton have Esingletons. By Distinctness some Esingletons, for instance the Esingleton of \( n \), are not Esingletons of diatoms consisting of \( n \) and an Esingleton. So if \( x \) is the fusion of all Esingletons of diatoms consisting of \( n \) and an Esingleton, and \( y \) is the fusion of all Esingletons, then \( x \) is a proper part of \( y \). Further, each Esingleton has as the Esingleton of the fusion of it and \( n \) exactly one Esingleton, by Functionality. Further, each Esingleton of a diatom consisting of \( n \) and one Esingleton is the Esingleton of exactly one such diatom, by Distinctness. So \( x \) is a proper part of \( y \), and each atom of \( y \) is the Esingleton of a diatom consisting of \( n \) and exactly one atom of \( x \), and each atom of \( x \) is such that exactly one atom of \( y \) is the Esingleton of a diatom consisting of \( n \) and exactly one atom of it. So by the Dedekind schema \( y \) is infinite, and since \( y \) exhaustively divides into atoms since it is a fusion of Esingletons, \( y \) is an infinite fusion of atoms. QED

Null Set: the null set is a set with no members.

Proof: \( n \) is the Null set. It is a class, by definition. It has no members, since no Esingleton is part of its fusion (from the atomicity of \( n \), and the axiom Null Set). From IR, there are an infinite number of atoms, so \( n \), being a single atom, is small. Since it is small and a class, it is a set. \( n \), therefore, is a set with no members.

Extensionality: No two classes have the same members; no class has the same members as the null set.

This needs to be modified, since this system treats the null set as a class, and of course it has the same members as itself. So instead I shall prove a modified version of Extensionality (the second clause may be redundant, but reduces possible ambiguity):

Extensionality*: No two classes have the same members; no class distinct from the null set has no members.

Proof: for classes to have the same members is for them to have the same Esingletons as parts (from the definition of membership). Since classes have no other parts besides \( n \) and Esingletons (from the definition of classes), and all have \( n \) as a part (also from the definition of classes), if two classes have the same Esingletons as parts they have all of their parts in common. This is impossible, by Uniqueness of Composition, so no two classes have the same members. The same holds in the special case where the classes have no Esingletons as parts, so there is at most one null class.
Pair Sets: if each of \( x \) and \( y \) is an individual or a set, then there exists a set of \( x \) and \( y \).

**Proof:** From Domain and the definitions of "set" and "individual", it follows that \( x \) and \( y \) have Esingletons, from Unrestricted Composition it follows that there is a fusion consisting of exactly the three atoms which are the Esingleton of \( x \), the Esingleton of \( y \), and \( n \): this fusion has \( x \) and \( y \) as members, and only \( x \) and \( y \) as members, since the Esingletons of \( x \) and \( y \) are the only Esingletons which it has as parts (\( n \) is not an Esingleton from Null Set). So this fusion is a class which has exactly \( x \) and \( y \) as its members (from the definition of "class" and "membership". This class has three parts, and from IR we know that there are infinitely many atoms, so this class is small, so, from the definition of "set", this class is a set. So there is a set which has exactly \( x \) and \( y \) as members.

**Aussonderung** (Separation): Given a set \( x \), and given some things, there is a set of all and only those of the given things that are members of \( x \).

**Proof:** If none of the given things are members of \( x \), then the required set is the null set. Otherwise \( x \) is a small class (from the definition of "set"); all members of \( x \) have Esingletons; so by Unrestricted Composition we have the fusion of all Esingletons of the members of \( x \) that are among the given things. This fusion fused with \( n \) is the class of all members of \( x \) that are among the given things (from definitions of "class" and "membership"). Since there are no more parts of this class than there are of \( x \) (since this class will be a part of \( x \), composed as it is entirely out of parts which are also parts of \( x \)), and since \( x \) is small, this class will be small also, so it will be a set (from the definition of "set").

Replacement: If there are some ordered pairs whereby each member of a class \( x \) is paired with exactly one member of a class \( y \), and if for each member of \( y \) there is a member of \( x \) that is paired with it, and if \( x \) is a set, then \( y \) is a set.

I will assume, with Lewis, that ordered pairs are defined in some standard set-theoretic way: or at least that one of the standard set-theoretic ways is an acceptable definition of one sort of ordered pairs.

This proof will employ the singular Replacement Schema, which is as follows:

If each atom of a thing \( x \) ... exactly one atom of a thing \( y \), and if for each atom of \( x \) there is an atom of \( y \) that ... it, and if \( x \) is small, then \( y \) is small. (Lewis p. 90).
Proof. The null class satisfies Replacement trivially, for the antecedent is necessarily false for any two classes one of which is the null class. For any two non-null classes which are such that they meet the condition specified by the antecedent of Replacement: consider the objects which are gained by mereologically subtracting n from each class (call these things X and Y). The replacement schema given allows substitution of any condition where the dots are, so in particular will allow for "is such that the object of which it is an Esingleton is paired by one of the given pairs with the object which has as its Esingleton", which yields:

If each atom of a thing x is such that the object of which it is an Esingleton is paired by one of the given pairs with the object which has as its Esingleton exactly one atom of a thing y, and if for each atom of x there is an atom of y that is such that the object of which it is an Esingleton is paired by one of the given pairs with the object which has as its Esingleton it, and if x is small, then y is small.

Distinctness will ensure that the Esingletons have unique members, and the definition of classes will ensure that each part of x and y are Esingletons. This instance of the Replacement schema, applied to X and Y, will ensure that if X is small, Y is small. Next it needs to be shown that if X+n is small then X is, and that if Y is small then Y+n is. (The alternative proof employs a more complicated substitution instance of the Replacement schemata). The sub-proofs go as follows: from the definition of small, if X+n is small, a proper part of it (i.e. X) will be too. The proof in the other direction is simple too: Y and n are few, n is small (since n is an atom and from IR there are infinitely many things), so if Y is small then Y and n are small and few, so by Hypothesis U their fusion is small. Alternatively, the "principle of the framework" connecting the infinite and finite to the large and small articulated on pp. 89-90 give immediately that any fusion of a small thing with a finite thing is small: since n is finite (it has one part, in fact), its fusion with any small thing is itself small. Small classes are sets (by definition of "set"), so if a class x is a set, i.e. a small class, and there is a class y which with the class x satisfies the antecedent of the Replacement axiom, y is a small class, and thus a set, also.

Choice: Suppose x is a class, and suppose there are some ordered pairs whereby each member of x is paired with at least one thing, and no two members of x are paired with the same thing. Then there is a class y such that each member of x is paired with exactly one member of y.

This proof will employ the First Choice Schema, which is as follows:
If there are some things, and each of them ... some things, and no two of them ... the same things, then there are some things such that each of the former things ... exactly one of the latter things.

**Proof.** Following Lewis, I will fill the blanks of the First Choice Schema with “is paired by one of the given pairs with”. (substituting “are” for “is” as needs be for grammatically...):

> If there are some things, and each of them is paired by one of the given pairs with some things, and no two of them are paired by [any] one of the given pairs with the same things, then there are some things such that each of the former things is paired by one of the given pairs with exactly one of the latter things.

This gives us that there are some things which are paired with the members of $x$ in the appropriate way, and that these things appear in pairs, and so have singletons. If they have singletons, then the fusion of the singletons of these things will yield a class (each singleton is a diatom consisting of an Esingleton and $n$, and any fusion consisting only of $n$ plus Esingletons is a class). So given the antecedent of Choice, there will be a class of things such that each member of $x$ is paired with exactly one member of it.

**Power Sets:** If $x$ is a set, there is a set of all subsets of $x$.

**Proof.** Since $x$ is a set, $x$ is small. Then the sub-classes of $x$ are also small (they are proper parts of $x$, after all), so they are sets, and have Esingletons. The fusion $y$ of those Esingletons with $n$ is the class of all subclasses of $x$ except for the null class. Fuse the Esingleton of $n$ with $y$, and the result $z$ is the class of all subclasses of $x$. By Hypothesis P, the parts of $x$ are few. Then by the Replacement Schema (plural version) the Esingletons of those parts are few. Since the Esingletons of the parts of $x$ which are classes are only some of the Esingletons of the parts of $x$, they must be few also. The fusion of these is small by the “principle of the framework” on p.91 of Lewis 1991. The fusion of these with $n$ and the Esingleton of $n$ will also therefore be small, (from the principle on pp 89-90). So this object $z$ will be small, and will be the class of all the subsets of $x$. So $z$ is a set of all subsets of $x$.

**Unions:** If $x$ is a set, there is a set of all members of members of $x$.

My proof here will be along the lines of the simpler one relying on Hypothesis U: I will not provide an analogue of the Burgess proof mentioned on p. 104 n3 of Lewis 1991. Note that there seems to be a lacunae in Lewis’s proof: he goes straight from (in his
proof) the fact that the singletons of the classes whose fusion is his $y$ are few to the result that those classes themselves are few. The lacunae can be filled with the invocation of the plural Replacement Schema (along with constraints on his singleton relation) in the manner analogous to my invocation of the plural Replacement Schema.

As well as Hypothesis U, this proof relies on the plural Replacement Schema, which is as follows:

Given some things, and given some other things (not necessarily different), if each of the former things ... exactly one of the latter things, and if for each of the latter things there is one of the former things that ... it, and if the former things are few, then the latter things are few.

Proof. If all members of $x$ are individuals or the null class, then the required set is the null set. Otherwise, consider those classes that are members of $x$. Let $y$ be their fusion; then $y$ is the class of members of members of $x$ (they are guaranteed to have a fusion by Unrestricted Composition). Each of these classes, since it is a member of something, must be small (from Domain). The class of them is part of the small class $x$ (and the parts of a small thing are few, from the principle of p.91), so this means that their Esingletons are few (since it has just been shown that the fusion of their Esingletons is a part of a small thing, therefore small). To show that they are few if their Esingletons are the plural Replacement Schema can be employed. Let the first group of things be the Esingletons of the classes in question, and the latter things the classes themselves, and let the blank be filled with "is an Esingleton of". The Esingletons will be Esingletons of exactly one of the classes each, by Functionality and Distinctness. The Esingletons are few, so by the instance of the plural Replacement Schema, the classes will be as well. The fusion of these classes (that is, $y$) is therefore a fusion of a few small things, hence it itself is small by Hypothesis U. Therefore $y$ is a set (from the definition of "set").

Infinity: There is a nesting with no greatest member, the union of which is a set.

This proof relies on the Second Choice Schema, which is as follows:

If nothing ... itself, and if whenever $x$ ... $y$ and $y$ ... $z$ then $x$ ... $z$, and if there are some things such that each of them ... another one of them, then also there are some of those things such that (1) among the latter things also, each one ... another one; (2) whenever $x$ and $y$ are two of the latter things, then either $x$ ...$y$ or $y$... $x$.

Proof. By Hypothesis I, there are some things such that each of them is a proper part of another of them (from the definition of "infinite" as it appears in Hypothesis I), and such that their fusion is a small fusion of atoms. Each of these things must be a small fusion of atoms (since each is a proper part of a small fusion of atoms). Replace each
atom by its Esingleton: that will not affect the smallness of the things and their fusion, nor will it affect the fact that the things are such that each is a proper part of another.

Now we can say that there are some sets such that each of them is properly included in another, and such that their union is a set. An instance of the Second Choice Schema with the blank filled by "is properly included in" is employed:

If nothing is properly included in itself, and if whenever \( x \) is properly included in \( y \) and \( y \) is properly included in \( z \) then \( x \) is properly included in \( z \), and if there are some things such that each of them is properly included in another one of them, then also there are some of those things such that (1) among the latter things also, each one is properly included in another one; (2) whenever \( x \) and \( y \) are two of the latter things, then either \( x \) is properly included in \( y \) or \( y \) is properly included in \( x \).

The antecedents are satisfied by the fusions of Esingletons, and so the desired nesting of fusions of Esingletons is provided: for there are some (the "latter things") which are such that each is included in another of them, and for any two, the first is properly included in the second of \( \text{vice versa} \). Since these are some of the original Esingleton fusions, they are small, and so fusing \( n \) to each will preserve their smallness, because the fusion of a small thing and a finite thing is itself small (from the principle on p 89-90). Fusing \( n \) to each of them yields small classes, so it yields sets. These sets are still nested by the proper-part nesting, since \( n \) is added to all of them and so leaves the extent to which any one overlaps any other unchanged. The fusion of these sets yields an object all of which except \( n \) is in the overlap with the small fusion of Esingletons originally produced by the original substitution of Esingletons. The object which is the fusion of these sets minus \( n \) is therefore small, and from the principle of p 89-90, the fusion of these sets is therefore small (for it is only one atom larger than a small thing). The fusion (and therefore the union) of all of these sets is a small class, so it is a set too. So this nesting with no greatest member is such that its union is a set.

**Deriving Fundierung from Foundation**

**Fundierung**: No class intersects each of its own members.

*Foundation Axiom of Esingleton*: Whenever there are some things, either at least one of them is not an object which is the fusion of at least one Esingleton with \( n \) and which contains no non-Esingletons besides \( n \) as parts, or one of them is such that all the objects which have Esingletons among its parts are not among those things.
Proof. By definition of class and membership, the only things with members are classes other than the null class: that is, the only things with members are fusions which have \( n \) as a part, have at least one Esingleton as a part, and have no non-Esingletons besides \( n \) as parts. So the first clause of the disjunction in the Foundation Axiom is satisfied if at least one of the things is a non-null class. The second is satisfied if there is a non-null class among the things, but all of its members (the objects whose Esingletons are parts of it, from the definition of membership) are not among the things. So Foundation is equivalent to the following claim in the language of classes:

Whenever there are some things, either at least one of them is memberless or there is something which is one of those things such that all of its members are not among those things.

Given this, Fundierung can be shown easily:

Suppose, for reductio, that there was a class \( x \) which intersected all of its members: that is, it overlaps each of the objects whose Esingletons are parts of it. It and all of its members are some things, so one of them is such that all of its members are not among those things. All of its members are among those things, so at least one of its members must be such that all of its members are not among those things. Call one such member \( y \). By hypothesis, none of \( y \)'s members are among the plurality consisting of \( x \) and its members. But by definition of \( x \) at least one of its members is a member of \( y \), since it intersects \( y \) because \( y \) is a member of \( x \). Reductio. So there is no such class as \( x \), so no class intersects all of its members.

This proof can be carried out without first translating Foundation into a claim about membership of course, but I present it in this form to make clearer Foundation’s affinity with standard formulations of the axiom of Foundation or Regularity in set theory.

The proofs for the second proposed system.

AXIOMS AND DEFINITIONS

1. Distinctness — no two objects share the same Esingleton
2. Functionality — nothing has more than one Esingleton
3. Domain — \( N \) has an Esingleton, any small fusion of atoms has an Esingleton, any fusion of a small fusion of Esingletons and \( N \) has an Esingleton and nothing else has a singleton.
4. **Atomicity** — all Esingletons are atoms.

5. **Foundation** — whenever there are some things, either at least one of them is not an object which is the fusion of at least one Esingleton with N and which contains no non-Esingletons besides parts of N as parts, or one of them is such that all the objects which have Esingletons among its parts are not among those things.

6. **Esingleton Distinctness** — every Esingleton is distinct from N.

Definitions: A **class** is any fusion which exhaustively divides into N and Esingletons (N counts as a class, and indeed is the null class, as n was). An object is a **member** of a class iff that object's Esingleton is a part of that class, and non non-class has a **member**. A **set** is any class which does not have a Large mereological difference from N. A **individual** is any small object.

**PROOFS**

**IR:** there is an infinite fusion of atoms.

This proof will employ the Dedekind Schema affirmed as a principle of the framework on p. 88-89 of Lewis, which is as follows:

If \( x \) is a proper part of \( y \), and if each atom of \( y \) ... exactly one atom of \( x \), and if each atom of \( x \) is such that exactly one atom of \( y \) ... it, then \( y \) is infinite.

**Proof.** By Domain, there are Esingletons, for instance, the Esingleton of N, and all fusions of N and exactly one Esingleton have Esingletons. By Distinctness some Esingletons, for instance the Esingleton of N, are not Esingletons of fusions of N and exactly one Esingleton. So if \( x \) is the fusion of all Esingletons of fusions of N and exactly one Esingleton, and \( y \) is the fusion of all Esingletons, then \( x \) is a proper part of \( y \). Further, each Esingleton has as the Esingleton of the fusion of it and N exactly one Esingleton, by Functionality. Further, each Esingleton of a fusion of N and exactly one Esingleton is the Esingleton of exactly one such fusion, by Distinctness. So \( x \) is a proper part of \( y \), and each atom of \( y \) is the Esingleton of a fusions of N and exactly one atom of \( x \), and each atom of \( x \) is such that exactly one atom of \( y \) is the Esingleton of a fusion of N and it. So by the Dedekind schema \( y \) is infinite, and since \( y \) exhaustively divides into atoms since it is a fusion of Esingletons, \( y \) is an infinite fusion of atoms. **QED**

**Null Set:** the null set is a set with no members.
Proof. $N$ is the null set. It is a class, by definition. It has no members, since by E singleton Distinctness it has no E singletons as parts. It of course does not have a Large mereological difference from itself, so it is a set. $N$, therefore, is a set with no members.

Extensionality: No two classes have the same members; no class has the same members as the null set.

This needs to be modified, since this system treats the null set as a class, and of course it has the same members as itself. So instead I shall prove a modified version of Extensionality (the second clause may be redundant, but reduces possible ambiguity):

Extensionality*: No two classes have the same members; no class distinct from the null set has no members.

Proof. for classes to have the same members is for them to have the same E singleton parts (from the definition of membership). Since classes exhaustively divide into $N$ and E singletons (from the definition of classes), and all have $N$ as a part (also from the definition of classes), if two classes have the same E singletons as parts they have all of their parts in common. This is impossible, by Uniqueness of Composition, so no two classes have the same members. The same holds in the special case where the classes have no E singletons as parts, so there is at most one null class.

Pair Sets: if each of $x$ and $y$ is an individual or a set, then there exists a set of $x$ and $y$.

Proof: From Domain and the definitions of “set” and “individual”, it follows that $x$ and $y$ have E singletons, from Unrestricted Composition it follows that there is a fusion exhaustively dividing into the E singleton of $x$, the E singleton of $y$, and $N$: this fusion has $x$ and $y$ as members, and only $x$ and $y$ as members, since the E singletons of $x$ and $y$ are the only E singletons which it has as parts ($N$ is not an E singleton from Atomicity). So this fusion is a class which has exactly $x$ and $y$ as its members (from the definition of “class” and “membership”). The part of this class distinct from $N$ consists of two atoms, and from IR we know that there are infinitely many atoms, so that part is small (if there is something infinite, then any finite thing is small from p.90 of Lewis), so, from the definition of “set”, this class is a set. So there is a set which has exactly $x$ and $y$ as members.
**Aussonderung** (Separation): Given a set $x$, and given some things, there is a set of all and only those of the given things that are members of $x$.

**Proof:** If none of the given things are members of $x$, then the required set is the null set. Otherwise $x$ is a class which is such that the fusion of its E-singlets is small (from the definition of “set”); all members of $x$ have E-singlets; so by Unrestricted Composition we have the fusion of all E-singlets of the members of $x$ that are among the given things. This fusion fused with $N$ is the class of all members of $x$ that are among the given things (from definitions of “class” and “membership”). Since there are no more parts of this class than there are of $x$ (since this class will be a part of $x$, composed as it is entirely out of parts which are also parts of $x$), and since the portion of $x$ distinct from $N$ is small, this class will be small also, so it will be a set (from the definition of “set”).

**Replacement:** If there are some ordered pairs whereby each member of a class $x$ is paired with exactly one member of a class $y$, and if for each member of $y$ there is a member of $x$ that is paired with it, and if $x$ is a set, then $y$ is a set.

I will assume, with Lewis, that ordered pairs are defined in some standard set-theoretic way: or at least that one of the standard set-theoretic ways is an acceptable definition of one sort of ordered pairs.

This proof will employ the singular Replacement Schema, which is as follows:

If each atom of a thing $x$ ... exactly one atom of a thing $y$, and if for each atom of $x$ there is an atom of $y$ that ... it, and if $x$ is small, then $y$ is small.

(Lewis p. 90).

**Proof.** The null class satisfies Replacement trivially, for the antecedent is necessarily false for any two classes one of which is the null class. For any two non-null classes which are such that they meet the condition specified by the antecedent of Replacement: consider the objects which are gained by mereologically subtracting $N$ from each class (call these things $X$ and $Y$). The replacement schema given allows substitution of any condition where the dots are, so in particular will allow for “is such that the object of which it is an E-singlet is paired by one of the given pairs with the object which has as its E-singlet”, which yields:

If each atom of a thing $x$ is such that the object of which it is an E-singlet is paired by one of the given pairs with the object which has as its E-singlet exactly one atom of a thing $y$, and if for each atom of $x$ there is an atom of $y$ that is such that the object of which it is an E-singlet is paired by one of the given pairs with the object which has as its E-singlet it, and if $x$ is small, then $y$ is small.
Distinctness will ensure that the Esingletons have unique members, and the definition of classes will ensure that each part of x and y are Esingletons. This instance of the Replacement schema, applied to X and Y, will ensure that if X is small, Y is small. The fusions X+N and Y+N will therefore be sets, since their mereological difference from N will be small. Y+N is the class y, so y is a set.

Choice:
The proof of choice is exactly as given above, except that the reason the fusion of the things paired with the members of x is a class is that fusions of singletons are classes, since singletons are fusions which exhaustively divide into Esingletons and N, and so a fusion of them will also, and any fusion which exhaustively divides into Esingletons and N is a class.

Power Sets: If x is a set, there is a set of all subsets of x.

Proof. Since x is a set, x minus N is small. Then the sub-classes of x also consist of something small fused with N (since the fusion of Esingletons which is the part of the subclasses disjoint from N are proper parts of x, after all), so they are sets, and have Esingletons. The fusion y of those Esingletons with N is the class of all subclasses of x except for the null class. Fuse the Esingleton of N with y, and the result z is the class of all subclasses of x. By Hypothesis P, the parts of x which are disjoint from N are few. Then by the Replacement Schema (plural version) the Esingletons of the fusions consisting of a part of x which is disjoint from N plus N itself will be few as well. The fusions consisting of a part of x which is disjoint from N fused with N itself are the non-null sub-classes of x, of course, since the only parts of x which are disjoint from N are fusions of Esingletons (from Atomicity and x's being a class as defined). The fusion of these is small by the "principle of the framework" on p.91 of Lewis 1991. The fusion of these with the Esingleton of N will also therefore be small, (from the principle on pp 89-90). The mereological addition of N to this object yields an object which has a small fusion of Esingletons as its difference from N, and that fusion of Esingletons contains the Esingletons of the non-null subclasses of x plus the Esingleton of N. This is therefore the object z. Since z-N is small, and z is a class, z is a set. So z is a set of all subsets of x.

Unions: If x is a set, there is a set of all members of members of x.

My proof here again will be along the lines of the simpler one relying on Hypothesis U: I will not provide an analogue of the Burgess proof mentioned on p. 104 n3 of Lewis 1991.
Proof. If all members of \( x \) are individuals or the null class, then the required set is the null set. Otherwise, consider those classes that are members of \( x \). Let \( y \) be their fusion; then \( y \) is the class of members of members of \( x \) (they are guaranteed to have a fusion by Unrestricted Composition). Each of these classes, since it is a member of something, must be a set, so must be a fusion of \( N \) with a small fusion of Esingletons (from definition of class and set). The classes which are members of \( x \) must only be few. To prove this, remember that the fusion of their Esingletons is small (since it is a part of the mereological difference between \( x \) and \( N \), since the classes in question are some of \( x \)' s members), and so those Esingletons will be few (from the principle that the parts of a small thing are few, from p. 91). Employment of the plural Replacement schema yields that if the Esingletons of the classes are few, then so are they, in the following manner. Insert in the blank of the Replacement schema "is an Esingleton of". By Distinctness, each Esingleton is the Esingleton of only one object, and all of the classes in question have an Esingleton among the Esingletons in question, so the antecedent of the instance of the Replacement schema is satisfied. So the instance tells us that if the Esingletons are few, the classes which they are the Esingletons of are few too. So those classes are few. So the fusion of the classes which are members of \( x \) is a fusion of a few things, each of which is \( N + \) a small fusion. So \( y \) is a fusion of \( N \) with a few small things, so the part of \( y \) disjoint from \( N \) is small, by Hypothesis U. \( y \) is therefore a class consisting of the fusion of \( N \) with a small thing. Therefore \( y \) is a set, from the definition of "set".

Infinity: There is a nesting with no greatest member, the union of which is a set.

Proof. By Hypothesis I, there are some things such that each of them is a proper part of another of them (from the definition of "infinite" as it appears in Hypothesis I), and such that their fusion is a small fusion of atoms (the first clause of this claim is also available from IR, of course). Each of these things must be a small fusion of atoms (since each is a proper part of a small fusion of atoms). Replace each atom by its Esingleton: that will not affect the smallness of the things and their fusion, nor will it affect the fact that the things are such that each is a proper part of another.

Now we can say that there are some sets such that each of them is properly included in another, and such that their union is a set. An instance of the Second Choice Schema with the blank filled by "is properly included in" is employed:

If nothing is properly included in itself, and if whenever \( x \) is properly included in \( y \) and \( y \) is properly included in \( z \) then \( x \) is properly included in \( z \), and if there are some things such that each of them is properly included in another one of them, then also there are some of those things such that (1) among the latter things also,
each one is properly included in another one; (2) whenever \( x \) and \( y \) are two of the latter things, then either \( x \) is properly included in \( y \) or \( y \) is properly included in \( x \).

The antecedents are satisfied by the fusions of Esingletons, and so the desired nesting of fusions of Esingletons is provided: for there are some (the "latter things") which are such that each is included in another of them, and for any two, the first is properly included in the second of vice versa. Since these are some of the original Esingleton fusions, they are small, and so each of them when fused with \( N \) will yield a set. The sets formed by fusing each of the nested fusion of Esingletons with \( N \) will still be nested by the proper-part nesting, since \( N \) is added to all of them and so leaves the extent of the mereological difference between any two unchanged. The fusion of all of these fusions of Esingletons with each other will yield a part of the fusion of Esingletons originally produced by the original substitution of Esingletons. The object which is the fusion of these Esingletons is therefore small. Its fusion with \( N \) then yields a class whose mereological difference from \( N \) is small, and so a set. This set is the fusion of all of the nested sets gained by fusing each of the nested fusions of Esingletons with \( N \), so it is the union of those nested sets. So this nesting with no greatest member is such that its union is a set.

**Fundierung:** No class intersects each of its own members.

**Foundation Axiom of Esingleton:** Whenever there are some things, either at least one of them is not an object which is the fusion of at least one Esingleton with \( N \) and which contains no non-Esingletons besides parts of \( N \) as parts, or one of them is such that all the objects which have Esingletons among its parts are not among those things.

**Proof.** By definition of class and membership, the only things with members are classes other than the null class: that is, the only things with members are fusions which have \( N \) as a part, have at least one Esingleton as a part, and have no non-Esingletons disjoint from \( N \) as parts (i.e. they contain no non-Esingletons besides parts of \( N \) as parts). So the first clause of the disjunction in Foundation is satisfied if at least one of the things is a non-null class. The second is satisfied if there is a non-null class among the things, but all of its members (the objects whose Esingletons are parts of it, from the definition of membership) are not among those things. So Foundation is equivalent to the following claim in the language of classes:
Whenever there are some things, either at least one of them is memberless or there is something which is one of those things such that all of its members are not among those things.

The remainder of the proof is as before.

Section 2: Proofs of Impredicative Comprehension

Lewis’s framework provides for a very powerful class theory — it is more powerful than Von Neumann-Bernays-Goedel class theory (e.g. that of Bernays 1937-1954, or of Goedel 1940). The class theory provided by Lewis is at least as powerful as Quine-Morse class theory (as it is called in Fraenkel, Bar-Hillel and Levy 1973). Other popular names for this class theory include Kelley-Morse, Quine-Kelley-Morse, or other combinations of the names of Quine, Kelley, Morse, Von Neumann, and probably Wang as well). In fact, on most measurements it is strictly more powerful, since it has plural quantification devices that quantify over proper classes, which gives second order resources greater than the class theory itself.

The distinctive feature of Quine-Morse class theory, according to Fraenkel, Bar-Hillel and Levy 1973 p. 138 is the Axiom of Impredicative Comprehension. They state the axiom (or rather an axiom schema for Impredicative Comprehension) as follows (p. 138):

There exists a class A which contains exactly those elements x which satisfy the condition $\mathfrak{B}(x)$, where $\mathfrak{B}(x)$ is any condition.

In this section I will prove that Lewis’s framework provides for this axiom schema for class theory, and will point out how the proof can be modified so that this principle applies to the proposals for producing class theories in chapter 7.

The proof is possible because the system of plural quantification which forms part of Lewis’s framework contains a powerful device of comprehension. Lewis says that a principle of plural ‘comprehension’ has “evident triviality” (Lewis 1991 p. 63) — though one might think that after the ungainly demise of naive set theory very powerful principles of comprehension can never again be evidently trivial. Nevertheless, we do seem to employ such a principle in our use of plurals in English, it has great intuitive support, and is it not known to lead to any contradictions. The principle, though not explicitly stated by Lewis, seems to be the following:

112 This observation is not original to me - I heard it first, I think, from A.P. Hazen. See Hazen 1996.
**Plural Comprehension:** For any condition C, if there is at least one C, then there are some things such that they are all and only the Cs.

(Lewis uses examples of “is a cat” and “is a set”, but appears to take the principle to hold generally). This is not the only important principle Lewis’s framework yields which is needed, but it (with Unrestricted Composition) is the most obviously relevant source of theoretical power. Let me now prove a version of Impredicative Comprehension in Lewis’s system (strictly speaking, impredicative comprehension should be rewritten for Lewis to read “set or class” instead of “class”, since for him the null set is not a class. Let me abbreviate “set or class” by “class” in the proof below):

**Proof.** For any condition C, either nothing satisfies C or at least one thing satisfies C (excluded middle). If nothing satisfies C, then the class which contains exactly those elements which satisfy the condition C is the null set (for there are no such elements). If at least one thing satisfies C, then there are some things (call them the Xs) such that they are all and only the Cs (from Plural Comprehension). Consider the condition “is one of the Xs and is an element”. Either nothing satisfies the condition or at least one thing does, by excluded middle. If nothing satisfies that condition, then the class of exactly those elements which meet condition C is again the null class. If at least one thing does, then there are some things such that they are Xs and are elements (Plural Comprehension). By definition of “elements”, all of those things have singletons. Since there is at least one of them there are therefore some things which are all and only the singletons of the elements which satisfy condition C (from Plural Comprehension). By Unrestricted Composition, there is a fusion of those things which are all and only the singletons of the elements which satisfy condition C. This fusion consists entirely of singletons, so it is a class. So for any condition, there is a class of all and only those elements which satisfy that condition. QED

The proof is very similar in the versions of the Framework suggested in chapter 7: the proof goes exactly as above, except that “null class” should be substituted for “null set” in the above, and the second last sentence should read “This fusion consists entirely of n and Esingletons, and so is a class”. In the system involving N, the second last sentence should be changed to read “This fusion exhaustively divides into N and Esingletons, and so it a class".
Appendix 2: Trouble for Lewis’s “Main Thesis”

I. The Problem

Lewis’s perhaps most central claim in Parts of Classes is his

Main Thesis: The parts of a class are all and only its subclasses.\(^{113}\)

Lewis takes one half of the Main Thesis to not need further support in the form of a proof of it from more self-evident considerations (though he does provide some reasons to make it plausible). The half of the Main Thesis not to stand in need of immediate proof is his

First Thesis: One class is a part of another iff the first is a subclass of the second.

The other half of the Main Thesis is the

Second Thesis: No class has any part that is not a class, and it is this half of the Main Thesis he claims stands in need of further argument. This further argument employs the following three premises:

Division Thesis: Reality divides exhaustively into individuals and classes.

Priority Thesis: No class is part of any individual.

Fusion Thesis: Any fusion of individuals is itself an individual.

One last thing needs to be mentioned: Lewis’s somewhat idiosyncratic definitions of classes and individuals. Later in (Lewis, 1991) Lewis redefines his central notions (pp 14-18), but this proof occurs before those redefinitions, and Lewis uses definitions stated on p 4 for the purposes of this proof. Lewis says

By ’classes’ I [i.e. Lewis] mean things that have members.
By ’individuals’ I [i.e. Lewis] mean things which are members, but do not themselves have members. Therefore [e.g.] there is no such class as the null class. (Lewis, 1991, p 4)

\(^{113}\) The statement of this thesis, as with the statements of the other theses, is taken directly from (Lewis 1991). The various theses can be found on pp 4-7. Essentially the same discussion can be found in (Lewis, 1993), on pp. 5-7.
Membership is taken to be primitive at this stage (see Lewis, 1991, p 16). Finally, Lewis uses the term “mixed fusion” to mean an object which divides exhaustively into individuals and classes: he holds that no mixed fusion is an individual (p 8). With these important definitions out of the way (and they do make a great deal of difference to what follows), let us look at Lewis’s argument from his four theses to the Second Thesis.

The proof of the Second Thesis offered by Lewis is quoted here in its entirety, with the steps I will argue are unsupported put in italics:

Proof Suppose the Second Thesis false: some class \( x \) has a part \( y \) that is not a class. If \( y \) is an individual, \( x \) has an individual as a part; if \( y \) is a mixed fusion of an individual and a class, then again \( x \) has an individual as a part; and by the Division Thesis those are the only possibilities. Let \( z \) be the fusion of all individuals that are a part of \( x \). Then \( z \) is an individual, by the Fusion Thesis. Now consider the difference \( x - z \), the residue that remains of \( x \) after \( z \) is removed. (It is the fusion of all parts of \( x \) that do not overlap \( z \).) Then \( x - z \) has no individuals as parts, so it is not an individual or a mixed fusion. By the Division Thesis, it must be a class. We now have that \( x \) is the fusion of class \( x - z \) with an individual \( z \). Since \( x - z \) is a part of \( x \), and not the whole of \( x \) (else there wouldn't have been any \( z \) to remove), we have that the class \( x - z \) is a proper part of the class \( x \). So, by the First Thesis, \( x - z \) must be a proper subclass of \( x \). Then we have \( v \), a member of \( x \) but not of \( x - z \). According to standard set theory, we then have \( u \), the class with \( v \) as its only member. By the First Thesis, \( u \) is a part of \( x \) but not of \( x - z \); by the Priority Thesis, \( u \) is not part of \( z \); so \( u \) has some proper part \( w \) that does not overlap \( z \). No individual is part of \( w \); so by the Division Thesis, \( w \) is a class. By the First Thesis, \( w \) is a proper subclass of \( u \). But \( u \), being one-membered, has no proper sub-class. This completes a reductio. (Lewis, 1991, pp 9-10, my italics).

Lewis’s error is to suppose that something which exhaustively divides into classes and which does not contain individuals can thereby be assumed to be a class. But this does not follow from his definitions or his principles: something which exhaustively divided into classes but which itself was not a class nor a member of any class would be neither individual nor mixed fusion, but would not conflict with the Division Thesis (as stated) by being some kind of thing which could not be broken down into classes or individuals, or both. Neither would it conflict with the Priority Thesis, since the putative object would not be an individual, nor necessarily a part of any individual. The First Thesis is silent on the point of whether classes contain things other than classes, and of course the Fusion Thesis is irrelevant in this case. The proof thus goes astray in three places — a class could have a non-class as a part without there being any \( z \) as defined, because it could have a part which exhaustively divided into classes but which was not itself a class. Furthermore even if one did assume that there was such a \( z \), there is no guarantee that \( x - z \) would be a class, because Lewis’s conditions do not ensure that
something is a class if it has no individuals as parts. Finally, were \( x - z \) a class after all, then \( w \), the proper part of \( u \) which is the overlap of \( u \) and \( x - z \) need only exhaustively divide into classes, and need not be a class itself.

Not only is the proof invalid as it is stated, it suffers the further defect (I claim) that it cannot be usefully repaired. For it to be usefully repaired, one of two things would need to be done: either i) a valid proof from the same postulates would have to be constructed, or ii) there would need to be a plausible postulate added (or one of the existing ones modified) so that a valid proof along the same general lines could be produced. But in the second case, it would furthermore need to be the case that the new postulate or modification would need to derive some of its plausibility from some other source than the Main Thesis or the Second Thesis, since otherwise the proof would be useless as a means of persuading one of the plausibility of the Second Thesis or the Main Thesis. I will provide a model which satisfies the four theses, in the form in which they are presented, but not the main thesis, and so dispose of the possibility of a remedy of the first sort. This model will illustrate the lacuna in the initial proof, and I will argue that no modification to the set of postulates which fills this lacuna will be independently motivated by intuition, and so no remedy of the second sort will usefully repair the argument either.

II. The Model

Suppose that singletons are atomic, and that classes have members in virtue of having singletons as parts. Suppose further that every class with more than one member has an atom \( a \) as a part. \( a \) itself is an individual, and so itself has a singleton (call this singleton \( A \)). The fusion of \( a \) and other individuals will itself be an individual, as usual. A fusion of more than one atom which exhaustively divides into singletons will not then be a class, and by the Priority thesis will not be an individual either. (Note that classes, so described, are analytically not classes according to Lewis's redefinition on p 16, but he is not entitled to employ his redefinition at this stage). Suppose that there are two individuals, \( b \) and \( c \), and let their corresponding singletons be \( B \) and \( C \). Take the class which is a fusion of \( a, B \) and \( C \) (call this class \( D \) — \( D \) is of course \( \{b, c\} \)). It has a part, \( B+C \), which is not a class, and neither is it an individual or mixed fusion, of course. Of course, a model of this sort which satisfied the usual axioms of class theory would also include an infinite number of other classes in the model (including \( A \) and its fusions with \( B \) and \( C \)), as well as the individuals which can be mereologically built up from \( a, b, \) and \( c \), but these can be taken as read. The First Thesis is satisfied, as all of the subclasses of \( D \) are parts thereof, and all the classes which are parts of \( D \) are its subclasses. No class is part of any individual, so Priority is satisfied, and of course Fusion need not be violated. Division is not violated, because even though there are entities like \( B+C \), which are neither individuals nor classes, they nevertheless can be divided exhaustively into classes.
Reality can still be divided exhaustively into individuals and classes — the fusion of all the classes plus the fusion of all the individuals is the whole of Reality, after all.

One might interpret the Division Thesis so that something stronger is intended — for instance, that not just Reality, but every object divides exhaustively into individuals and classes\textsuperscript{114}. This still will not help, since in the model all of the atoms are individuals or classes (a is an individual, and the rest are normal individual atoms or singletons). It would mean that the first step of Lewis’s proof which I italicised would work, however. Of course, one could always propose a thesis to replace the Division thesis which would do the required work — such as the

\textit{Disjunctive Thesis:}\textsuperscript{115} each thing is an individual, or a class, or has both individuals and classes as parts.

As an exegetical note, this might well have been what Lewis was trying to capture all along with the Division Thesis, for towards the end of Lewis 1991 he states the Division Thesis as follows:

\begin{quote}
Reality divides exhaustively into individuals and classes; in other words, everything is an individual, a class, or a mixed fusion of individual and class. (Lewis 1991, p 99)
\end{quote}

I have already shown that the second statement is not the originally stated Division Thesis “in other words”, for there is a model where Reality divides exhaustively into individuals and classes, but the putative restatement does not hold. Nevertheless, it may have been what Lewis intended, and if so Lewis’s intended argument goes through. The Disjunctive Thesis would disallow the model, since B+C fails to satisfy it. But it appears terribly \textit{ad hoc}, and could hardly make the Second Thesis more plausible than it is without this argument. Lewis does offer an argument for the Division thesis which one might think can be adapted to support the Disjunctive Thesis, and so defend it from the \textit{ad hoc} charge. It is as follows:

\begin{quote}
All I can say to defend the Division Thesis, and it is weak, is that as yet we have no idea of any third sort of thing that is neither individual nor class nor mixture of the two. (Lewis 1991, p 8)
\end{quote}

\textsuperscript{114} It is clear that the word “Reality” for Lewis is meant to refer to the fusion of everything rather than to everything distributively: see Lewis, 1991, p 8. However, this interpretation of the Division thesis as the stronger division thesis might have been what Lewis had in mind in this argument, even if this is not reflected unambiguously in the formal statement of the principle. Evidence for this includes not only the informal gloss on the Division Thesis on p. 7, but also the discussion of the “Lasso Hypothesis” on p. 43. (My thanks to Lewis for bringing the Lasso Hypothesis discussion to my attention in this context).

\textsuperscript{115} Unlike the other Theses discussed, the Disjunctive Thesis is not to be found under that name in Lewis’s work.
I do not see that this is even a useful argument for the Division Thesis, (even when it is only supposed to be weak). To take this to be intuitive support for the Division Thesis is to make the same mistake Lewis (in my opinion rightly) accuses other theorists about classes of making. Lewis, when discussing the view that we know that classes are outside space and time, says

We go much too fast from not knowing whether they are [in space and time] to thinking we know that they are not, just as the conjurer’s dupes go too fast from not seeing the stooge’s head to thinking they see that the stooge is headless. (Lewis, 1991, p 33)

Similarly, it seems, to defend the Division Thesis (or the Disjunctive Thesis) by the consideration he adduced is to go too fast from having no idea whether there is any third kind of thing to claiming that there is not. It has a further problem when used to support the Tripartite Thesis, in that the model does give us some idea about what the “third sort of thing that is neither individual nor class nor mixture of the two” would be — a mixture of classes which was not itself a class. Perhaps we have no idea whether or not fusing classes always yields classes before beginning metaphysical theorising — but it is hardly plausible that it is intuitive that it cannot be that fusing classes may yield non-classes. So the Disjunctive Thesis could only save the argument by introducing a premise as unintuitive as the Second Thesis (and perhaps less intuitive), which would destroy the dialectical usefulness of the argument. This problem will arise for any other such repair to the argument: any principle strong enough to rule out unwanted cases will itself have as little or less intuitive support than the Second Thesis or Main Thesis themselves. Useful repair does not seem possible.

III. The Remedy

The proof cannot be usefully repaired, I claim. So what is to be done? Is the well-named Main Thesis of *Parts of Classes* in jeopardy, and does this spell trouble for the theory of classes outlined? Not really. For Lewis can retain the Main Thesis without relying on a proof for the Second Thesis, and just take the Main Thesis as the starting point for his theory. If a justification is needed for it, Lewis can justify it through the fruits of its labours: the theory of classes in *Parts of Classes* is in many ways an attractive theory, and presumably Lewis would hold that we are justified in believing it on the basis of the benefits it brings in explaining set theory. If he is right that it is a good and useful theory, then Quinean holistic considerations would justify us in believing the Main Thesis. And if neither it nor any close cousin is a good and attractive theory, the contentious argument
for the Second Thesis would not save the Main Thesis anyway. So all Lewis ought to do is abandon the proof while retaining the conclusion.\textsuperscript{116}

\textsuperscript{116} Thanks to Mark Colyvan and James Chase for helpful criticism, and thanks to David Lewis for helpful discussion and criticism of an earlier draft of this paper.
References


Boolos, George. 1984. “To Be is To Be the Value of A Variable (or To Be Some Values of Some Variables)”, Journal of Philosophy, 81: 430-439


Devitt, Michael. 1980. "'Ostrich Nominalism' or 'Mirage Realism'?", Pacific Philosophical Quarterly, 61: 433-439


Lycan, W. 1993: “Armstrong’s New Combinatorialist Theory of Modality” in Bacon, Campbell and Reinhardt 1993, pp 3-17


