

Université de Montréal  
Faculté d'Arts et Science

Ce mémoire intitulé:  
Lewis' Theory of Counterfactuals and Essentialism

Présenté par:  
Ian Lippiatt

A été évalué par un jury compose des personnes suivants:

Jean-Pierre Marquis  
(président-rapporteur)

François Lepage  
(directeur de recherche)

Frédéric Bouchard  
(membre du jury)

La logique contemporaine a connu de nombreux développements au cours de la seconde moitié du siècle dernier. Le plus sensationnel est celui de la logique modale et de sa sémantique des mondes possibles (SMP) dû à Saul Kripke dans les années soixante. C'est dans ce cadre que David Lewis exposera sa sémantique des contrefactuels (SCF). Celle-ci constitue une véritable excroissance de l'architecture kripkéenne. Mais sur quoi finalement repose l'architecture kripkéenne elle-même ? Il semble bien que la réponse soit celle d'une ontologie raffinée ultimement basée sur la notion de mondes possible. Ce mémoire comporte quatre objectifs. Dans un premier temps, nous allons étudier ce qui distingue les contrefactuels des autres conditionnels et faire un survol historique de la littérature concernant les contrefactuels et leur application dans différents champs du savoir comme la philosophie des sciences et l'informatique. Dans un deuxième temps, nous ferons un exposé systématique de la théorie de Lewis telle qu'elle est exposée dans son ouvrage *Counterfactuals*. Finalement, nous allons explorer la fondation métaphysique des mondes possible de David Lewis dans son conception de Réalisme Modal.

Mots Clés: Conditionnels, Contrefactuels, Mondes Possible, Sémantique, Logique.

Modern logic since the end of the Second World War has undergone many developments. Two of the most interesting of these are the Kripkian Possible World Semantics and Lewis' system of Counterfactuals. The first was developed by Saul Kripke in the 1950s and 1960s and the second was developed by David Lewis in the 1970s. In some senses we can say that Lewis' system of counterfactuals or Counter Factual Semantics (CFS) is built on top of the architecture which Kripke created with his Possible Worlds Semantics (PWS). But, what is the Kripkian Possible World Semantics itself built on? The answer it seems is very finely tuned ontology founded on the notion of possible worlds. This paper will attempt to do the following. First, attempt to draw a distinction between on the one hand conditionals and the other counterfactuals and at the same time attempt to look at some of the historical literature surrounding counterfactuals and their application in various fields like the philosophy of science. Second, recapitulate Lewis' system of counterfactual semantics as developed primarily in Lewis' book *Counterfactuals*. Finally this paper will attempt to explore the metaphysical foundations of the possible worlds account argued for by David Lewis in his conception of Modal Realism.

Key Words: Conditional, Counterfactual, Possible Worlds, Semantics, Logic

*This paper is dedicated to my teachers:*

*To James Ring, for showing me that doors that I thought were closed to me were actually open to me.*

*To Ray Jennings, for being patient with me.*

*To Francois Lepage, for tolerating my nonsense.*

**Table of Contents:**

Introduction	p 1
Chapter 1: Conditionals	p 7
Chapter 2: Lewis's System	p 30
Chapter 3: The Nature of Other Worlds, Concrete or Abstract?	p 75
Conclusion	p 107
Citations	p 110
Bibliography	p 112



## Introduction

Possible worlds and the subsequent analysis of counterfactuals have proven to be very useful and have shaped the direction that modern analytic philosophy has taken in the last half century. Problems like the nature of modality and causation which were once considered among the most intractable, have become more accessible with recourse to strategies provided by these two developments. Possible world semantics allowed for the explanation of deeply problematic issues with regards to modal claims. The basic idea of possible world semantics is that notions of validity, soundness and completeness can be defined for a logic of modality in terms of models constructed from sets of alternative possible worlds. In the 1950s important results were reached by this method most notably by Saul Kripke. The most interesting of these results are the proofs which depend on the relation of accessibility between worlds, one world is said to be accessible from another if and only if the statements of the former are possible in or from the point of view of the latter. The different systems of modal logic are appropriate depending on how this relation of accessibility is construed in terms of the familiar logical properties of reflexivity, transitivity and symmetry.

The formal results of possible worlds semantics are well established and accepted however its application to areas of metaphysics is not without controversy, at issue is the metaphysical reality of different sorts of possibility and necessity. Additionally accounts of the logic of time and morality have been given analogous treatments to that of modal logic.

One of the most interesting and controversial aspects of this account of modality by means of possible world semantics is the nature of the possible worlds themselves. Typically, mathematical logicians need not usually be concerned with this, many sets of objects both real and fictional can be said to have the properties necessary of sets 'of worlds' for whatever purposes they may have. If the metaphysically robust issues of modality (for example whether there are more possible colors than we ever see) are to be tackled in terms of possible worlds then the nature of possible worlds must be taken seriously. Two positions have emerged with regards to the nature of these possible worlds. The first, could be called 'Modal Eliminativism', this position holds that there are no serious metaphysical roles for possible worlds. The second of these could be called 'Modal Realism', this position holds that possible worlds are concrete totalities just like the real world. The most famous proponent of the second view is the American philosopher David K Lewis. Lewis has argued that his account makes possible reductions of modality (both logical and causal) to more concrete notions, which are more easily dealt with. Likewise, propositions can be dealt with in a similar way with recourse to these concrete totalities. Others have suggested that we need not be forced to pick between what they say are equally radical positions. They have argued that there are non actual possible worlds; however, they are not like the actual world rather they are more like sets of propositions or other abstract entities.

Debate about possible worlds also extends to the related treatment of subjunctive conditionals and more specifically to counterfactuals. Generally counterfactuals have been thought to be of the following form, 'if it were the case that  $\theta$  then it would be the case that  $\psi$ '. What is required for this statement to be true is that among the possible worlds where  $\theta$  is true some world in which  $\psi$  is true be more similar, in the relevant respects, to the actual world than



any world in which  $\psi$  is false. Counterfactuals are sometimes called contrary to fact conditionals because they are understood to be counter to the facts of the actual world. Conditionals themselves are hypothetical statements of the form ‘if  $\theta$  then  $\psi$ ’ or ‘ $\psi$  if  $p$ ’,  $\theta$  is known as the antecedent and  $\psi$  is known as the consequent. Conditionals have classically been divided along the lines of indicative and subjunctive. Indicatives have typically looked something like the following, ‘if Oswald did not kill Kennedy, then someone else did.’ Whereas on the other hand subjunctives have looked something like the following, ‘if Oswald had not killed Kennedy, then someone else would have.’ Most subjunctives and all indicatives are ‘open’, because they presuppose nothing about the antecedent. Unlike ‘if John had won, he’d be rich’ or ‘if John should have won, he would be rich’ or ‘if John won, he is rich’ all of which imply that John did not win. The difference between counterfactuals and open subjunctives is less pertinent than the difference between subjunctives and indicatives.

Counterfactuals presuppose rather than assert the falsity of their antecedents. The following example shows this quite clearly. ‘If Clinton had been president, he would have been famous’. This seems inappropriate and out of place but not false because Clinton was in fact president. Again, the distinction between counterfactuals and open subjunctive is less important than the difference between subjunctive and indicative. The indicative statement about Kennedy is true yet the subjunctive is probably false. If we replace ‘someone’ with ‘no one’ the truth-values reverse.

Another interesting aspect of counterfactuals is that they may not be truth functional. A truth functional compound is one whose truth value is completely determined in every possible instance by the truth values of its components. For example ‘the president is a grandfather’ and

‘the president is childless’ logically entails the falsity of ‘the president is a grandfather and the president is childless.’ On the other hand ‘if the president were a grandfather then the president would be childless’ is false but ‘if the president were a grandfather then the president would be a father’ is true even if its components are false. This indicates that the truth of counterfactuals may at least in part be due to the contents of its components. This property is shared by both indicative and subjunctive conditionals. In contrast, the material conditional, which takes the form  $\theta \supset \psi$ , and is defined by stipulation as true when  $\theta$  is false or  $\psi$  is true and generally held to be truth functional.

The fact that counterfactuals are neither strict nor material conditionals created the ‘problem of counterfactual conditionals’ raised by Roderick Chisholm and Nelson Goodman. The problem centers around establishing what the truth conditions of counterfactuals are and how they are determined by their components. The approach taken by these two philosophers became known as the ‘metalinguistic approach’ and is very close to the deductive nomological model of explanation. A counterfactual is true when its antecedent is conjoined with the laws of nature and a statement of background conditions and it thus logically entails its consequent. The following example shows this approach, if we take the statement ‘if the switch had been flipped the light would be on’. This statement is true because the statement plus the relevant laws of electricity and the fact of the arrangements of the circuit entail that the light is on. The problem is to specify which facts are to be ‘fixed’ for any given counterfactual context. Here, background conditions cannot include the denial of the antecedent or the consequent, even though they are true, nor anything else that would not be true if the antecedent were.

The most interesting approach to counterfactuals comes from David Lewis, which can be seen as an extension of the possible worlds semantics developed for modal logic which says that a counterfactual is true when its consequent is true in the nearest possible world in which the antecedent is true. If we take the example of the light switch, that particular counterfactual is true when a world in which the switch was flipped and the light is on is closer to the actual world than one in which the switch was flipped but the light is not on. The main problem then is to specify which world is nearest for any given counterfactual and context. The difference between indicative and subjunctive conditionals can be accounted for then in terms of either a different set of background conditions or a different measure of nearness.

Counterfactuals turn out to be present in a number of interesting philosophical and scientific contexts. For example in distinguishing between laws and accidental generalization. Laws tend to be counterfactually functional in the sense that they imply their own proper counterfactual. An example illustrates this quite well, if we take the law 'All copper conducts electricity' then this implies that had a particular piece of metal been made of copper it would have conducted electricity. On the other hand if we take the accidental generalization 'everything in my pocket conducts' it doesn't seem to imply that if an object had been in my pocket it would be one that conducted electricity. Of course the same can be said to apply for all dispositional properties like solubility and fragility. Counterfactuals pop up in other areas, for example it has been suggested that cause and effect can be understood in terms of counterfactuals in such a way that one event is the cause of another if and only if the latter would not have occurred if the former had not. For example 'if the bomb had not gone off then the bridge would not have collapsed.' Additionally counterfactuals have reared their heads in the debates concerning free will. An action is thought to be free only if the agent could have or

would have done otherwise if he had wanted to. Similarly with regards to moral philosophy an action is considered right only if a rational and informed agent would have chosen it. Finally in the philosophy of mind, a person can be said to be in a particular mental state if and only if he would behave in certain ways given certain stimuli. These are just some of the areas counterfactuals have proven to be useful in various fields of study.

Given all of this, we will attempt to do the three following things. First, we will explore the distinction between indicative and counterfactual statements and more broadly the nature of conditionals with regards to their truth values and their place in fact stating discourse. Second, we will give a detailed account of David Lewis's System of counterfactuals. Finally we will evaluate the metaphysical underpinnings of the possible worlds account paying special attention to its relationship to Lewis's system of counterfactuals.

## Chapter 1: Conditionals

*If the dog is in the yard then the mailman won't deliver the mail. If the dog had been in the yard then the mailman would not have delivered the mail.* These two sentences, or ones just like them, represent typical everyday utterances and expressions heard and understood by normal individuals. Their meanings are usually (at least by the non-philosophically inclined) not questioned and are used to convey information by speakers to those who hear them. In most instances, the speakers of such sentences feel that their meaning has been made clear and that they have communicated information, to those who hear them, in such a way as to achieve the purposes they had in mind when they chose to utter them. But, what exactly are they? The answer is far from clear and presents one of the most complex problems in the philosophy of language. We tend to identify them both as *conditional sentences*. A conditional sentence, in its most basic form a sentence with two parts or clauses. The first part, is what is called the 'antecedent' and is prefaced by the word 'if' and the second part, is known as the 'consequent' and is usually (though not always) preceded by the word 'then'. We tend to want to say that the consequent *follows* (logically, inferentially, causally) from the antecedent in such a way that it is impossible to hold the antecedent without thereby coming to hold the consequent. The first of the sentences seems, at least at first glance by any native speaker, to express a state of affairs, such that, the presence of the dog shall impede the delivery of the mail. In this way a condition is set by the presence of the dog in the yard, where a second condition seems to result, that is, the non delivery of the mail by the mailman. This type of sentence is thus known as an *indicative*

*conditional*. The second sentence, seems to be different however, it does not seem to report on an *actual* state of affairs but rather something much more hypothetical or speculative. The second sentence seems to report on an event that did not occur, the dog was not in the yard, but, *had he been*, the mailman *would not have* delivered the mail. The first condition (the dog being in the yard) wasn't met, thus the second condition (the non delivery of the mail) was not met either. This type of sentence is known as a *subjunctive conditional*. Although this division between these two types of sentences has intuitive appeal and seems to be well supported by both normal use by speakers and analysis of grammar, the story of these two sentences and their relationship is far from decided. We will attempt to evaluate some of the pertinent issues with regards to conditionals. Namely, we will take up the three following issues; first we will study the debate concerning the indicative and subjunctive distinction, secondly we will pursue the question of whether conditionals are propositions, and thirdly we will analyze the issue of whether we should see conditionals as part of fact stating discourse.

It should be stressed from the outset that very little consensus has emerged with regards to conditional sentences. Though it may be hard to believe that two short words like *if* and *then* joined together in sentences used everyday in normal discourse should cause such consternation. There is very little agreement about both details and fundamentals and nearly everything about these sentences is in fact at issue. Whether or not a unified theory of conditionals is even possible has been called into question and some philosophers of language have indeed argued that the attempt to find a unified theory of conditionals is hopelessly misguided because there may in fact be irreconcilable and wholly different kinds of conditionals.

### Subjunctive or Indicative?

There has been a traditional inclination to identify conditionals as falling into one of two categories, the indicatives and the subjunctives. Yet, it seems at times as though this distinction is incapable of truly capturing the whole story. Questions regarding substantive aspects related to the temporality of the sentences in each category can be raised. This is illustrated as follows. (1) *If Hamas didn't do it, Hizbolla did.* (2) *If Hamas doesn't do it, Hizbolla will.* These are both indicative conditionals, however (1) seems to be directed towards the past while (2) seems to be directed towards the future. In the first we seem to be speaking of events that have already occurred while in the second we speak of events which have yet to occur. The same situation applies to the subjunctive sentences, illustrated as follows. (3) *If Hamas had not done it, Hizbolla would have.* (4) *If Hamas were not to do it Hizbolla would do it.* Again, it seems as though both are subjunctive conditionals however, the first is directed towards the past while the second is directed towards the future. Although in contrast to the indicatives, the events mentioned in both (3) and (4) have not yet occurred or may never occur.

There is a tendency to look at indicatives as 'indicating' or pointing towards situations or states of affairs. On the other hand the subjunctives have been classically seen as dealing with something more hypothetical and indeed 'hypothetical-after-the-fact'. As the name suggests, it is a 'subjuncture' or breaking off of a path from another. This is at least the sense the distinction is supposed to have intuitively. What role our intuitions about these two types of conditionals should play in the discussion is far from clear. The suggestion has been made that it is indeed our intuitions that we should look to, as a guiding light when dealing with fundamental issues in

grammar like this. However others have suggested that it is precisely our blind pursuit of intuition that has led us so far astray.

Two views have generally emerged regarding the nature of these sentences. There are of course the Monists, who have held that it makes little or no sense to maintain any distinction between indicative and subjunctive conditionals, rather we should collapse them into one super class of conditional. On the other side then, there have been the Dualists, who have held that a distinction between indicative and subjunctive sentences does in fact make sense and is useful in the study of the issues at stake here. We will attempt to explore the merits of both positions reserving judgment until later.

In *On Conditionals* (1995) Dorothy Edgington argues that the “traditional distinction is less between two species of a genus than between two genera requiring separate treatment.”<sup>1</sup> Furthermore it seems as though whether we chose to call a sentences indicative or subjunctive, hinges on the occasion it was uttered; before, during or after the event in question or perhaps even with regards to known or unknown information . Authors like V.H Dudman have argued that the subjunctive/indicative distinction is horribly misguided, he claims that “English boasts nothing that merits the term ‘subjunctive’ and therefore nothing that merits ‘indicative’. English is without distinction of mood. Those many logicians and philosophers who take the indicative/subjunctive dichotomy as fundamental go wrong at their very first step”<sup>2</sup> In contrast to this Ernest Adams argues that there are indeed “plausible reasons for thinking that subjunctive and indicative conditionals differ in important respects, though they have interesting logical interrelations.”<sup>3</sup> As suggested before little consensus has emerged in regards to these issues.



The questions surrounding the subjunctive or ‘contrary to fact’ conditionals (as they are sometimes called) came to light first as a problem in the philosophy of science, it was hopped by the Logical Empiricists that they might be used as a means of more adequately founding scientific laws in such a way that the logical system, of which Frege had recently developed, could be used in the sciences. As all students of modern logic know, Frege’s classical interpretation of the conditional held that the conditional is false only when the antecedent is true and the consequent is false. This is what has been called ‘horseshoe’ and is usually written using the following notation  $\theta \supset \psi$  and is read as ‘if  $\theta$  then  $\psi$ ’, ‘ $\psi$  if  $\theta$ ’, and ‘ $\theta$  implies  $\psi$ ’. However, problems with this interpretation quickly became apparent in scientific contexts with regards to dispositional properties like fragile and soluble. It seems if we consider sentences like; *If the vase drops, it will break* (ascribing the property of fragility to the vase) and *If the sugar cube is submerged in water, it will dissolve* (ascribing the property of solubility to the sugar cube). Questions emerged about whether these two objects actually have these properties if they fail to undergo these changes, that is, if the vase is never dropped or the sugar cube is never submerged. We seem to want to say, that they do, however we have very little good reason to think so without actually undergoing the process in question. Further problems emerge when we wish to generalize these statements from individual vases and sugar cubes to *all* or even *most* vases and sugar cubes we have dealings with. Carnap explored some of these issues in *Testability and Meaning* (1936). Nelson Goodman argued that the problem of counterfactuals is not merely a grammatical exercise, rather if the problem of counterfactual conditionals can’t be adequately resolved then we have little hope of achieving a satisfying philosophy of science. Indeed, what was at stake according to Goodman was a definition of a scientific law and a proper theory of confirmation, along with an explanation of the nature of dispositional properties. Authors like

Kit Fine have interpreted Goodman's account of counterfactuals in the following way, he suggests that the "counterfactual  $\theta > \square \psi$  is true if  $\psi$  is a logical consequence of  $\theta$  and some other sentences which are either implicit from the context or cotenable with  $\theta$ . Cotenability is in need of explanation, but for a wide range of counterfactuals the cotenables can consist of laws of nature and certain auxiliary conditions."<sup>4</sup> Fine has argued that we can formalize this in the following way, we let " $|\theta|$ " be the truth set of the sentence  $\theta$  i.e. the set of worlds in which it is true. Then we can say that the counterfactual  $\theta > \square \psi$  is true at world  $i$  if there are sets of worlds which are cotenable with  $|\theta|$  and such that  $|\psi|$  contains their intersection with  $|\theta|$ . Here cotenability is a primitive relation between sets of worlds and dependent upon the given world  $i$ ."<sup>5</sup> This seems to give Goodman's account theoretical foundation in the possible worlds semantic; whether it really illuminates the issue at all or merely lifts and shifts the problem somewhere else, I leave to the reader to decide

There may indeed be very good reason for not collapsing the indicative and subjunctive classes of conditionals into one overarching super class of conditional. If we turn our attention to the two following sentences: (1) *If John Wilkes Booth didn't kill Abraham Lincoln, someone else did*, and (2) *If John Wilkes Booth hadn't killed Abraham Lincoln, someone else would have*. The first is obviously an indicative where the second is obviously a subjunctive. If we consider two further sentences; (O) *John Wilkes Booth didn't kill Abraham Lincoln*, and (S) *someone else killed Abraham Lincoln*, along with the following sentence frames

(f1) *if it is the case that-----it is the case that-----.*

and

(f2) *if it were the case that-----it would be the case that-----.*

Clearly we can take (O) and (S) and fill in the sentence frames in their respective antecedent and consequent spaces. The result accordingly yields two distinct sentences, with distinct *meaning* in such a way as to make it possible to accept the filled in sentence frame (f1) while rejecting sentence frame (f2). However, even if this is convincing, the conclusion that they are vastly different sentences expressing different truths, can in fact be resisted. The argument for different accounts of ‘if’ may in fact misrepresent the situation. Consider the following four sentences

If he kills the president

If he were to kill the president

If he killed the president

If he had killed the president

Each of these sentences may in fact represent the same conditional thought. This some have suggested, shows that the categories of indicative and subjunctive with regards to conditional sentences are highly problematic and a sign that our accounts of them still need some work. Edgington seems to claim that a true counterfactual can have a true antecedent and a true consequent, however in her words counterfactuals are for “talking about unrealized possibility” and further “we use them when we think the antecedent is false.”<sup>6</sup> In contrast to Edgington, Dudman and Bennett have argued that ‘had been’ and ‘would’ concern tense not mood. They seem to want to make the problem a temporal one rather than predominately a modal one. They

find that it is a mistake to think of these counterfactuals as expressing anything like possibility in their meanings.

With regards to the distinction between indicative and subjunctive conditionals, perhaps Edgington is correct when she says that the difference between them is maybe “more like the difference between mature cheddar and freshly made cheddar” rather “than the difference between chalk and cheese.”<sup>7</sup> Clearly, indicative and subjunctive conditionals are related, indeed they seem necessarily intertwined with each other in the sense that any indicative (or subjective) conditional automatically seems to suggest its sibling. However, the drive to merge them into a super class of conditional is not appropriate, given that it seems to shed no more light on the issues at play here. In some instances, it seems as though the push for a monistic conception of conditionals confuses rather than illuminates. Consequently a dualist approach (at least for now) seems wiser. For better or worse, the distinction between indicative and subjunctive conditional sentences holds. We now turn to the separate issue of whether or not conditionals should be seen as propositions

### **Are Conditionals Propositions?**

If we look at a sentence like the following, from a famous children’s song ‘if the bough breaks the cradle will fall.’ Can we say that it is a proposition? Clearly it has two parts, the antecedent ‘the bow breaks’ and the consequent ‘the cradle falls’ and each of them is a proposition, but it’s not clear (at least not definitively) whether or not the resulting conditional is a proposition. A quick word on propositions. There is fierce debate as to the nature and indeed

the existence of propositions. Classically, we have understood a proposition to be something which puts forward a claim to describe a state of affairs. If we take the example 'my cat is white' the statement sets up conditions that are either fulfilled or not. In the case of our example the following conditions are laid out 1) that there is a person referred to by 'my' (the speaker), 2) that there is a creature known as a cat, 3) that this creature and the speaker can stand in a relationship where the speaker can indeed own the cat, and 4) that the cat can have the property of being white. A proposition then, has conditions which render it as true or false, what is important is that our bit of language ('my cat is white') can in fact carry a truth value based on the conditions we laid out. Our sentence is a proposition because it lays out conditions which either succeed or fail whatever the case may be. As it happens for our example, it is indeed a proposition but it is false, my cat is actually gray. Other examples of false propositions are things like 'Shanghai is in Japan', 'Napoleon died in Paris' and, 'Andy Kaufman was from the moon'. Clearly not all bits of language are propositions. Something like 'Shut the door!' which is an order or command is not a proposition and it cannot carry a truth value because it doesn't set up any conditions to be fulfilled. A related question, concerns whether or not moral judgments are propositions or not (and thus lack objective truth values). To hold that moral judgments are not propositions is the position of Moral Expressiveness, that is to say that the meaning of the expression of a moral judgment (for example 'eliminating abortion would be good') is merely to express the attitude of the speaker and consequently statements of moral judgment have no truth values. I should stress that the Moral Expressivist holds this position in service of a metaphysical thesis; that there are no objective moral facts or properties, whereas those who take the position that conditionals are not propositions (and thus lack objective truth values) do so in service of no such metaphysical thesis. I shall have no more to say about this and raise it

only as an interesting parallel. Everything I have just said is of course a vast oversimplification, however, we need not concern ourselves with the questions of what propositions are and whether they exist or not (at least not too much), all we must say is that there are bits of language which are propositions, they can have truth values (true or false as the case may be) and these propositions can be connected and the resulting bit of language is, or is not, itself a proposition.

For our purposes we can ask the question ‘when we fill in the gaps in the conditional, with propositions, is the resulting sentence a proposition?’. The answer to this question is far from clear. The related question arises whether the material conditional, represented classically by the  $\supset$  in if-then conditional sentences is more like  $\vee$  and  $\wedge$ , or is it more like the modal operators,  $\square$  and  $\diamond$ . A quick illustration, clearly (1) ‘we had lobsters last night’ and (2) ‘we had oysters last night’ are both propositions, they can have truth values because again they set up conditions which are either fulfilled or not. Clearly we can take these two proposition and combine them with connective ‘and’ to get ‘we had lobsters last night and we had oysters last night’ for simplicity we can just say (3) ‘we had lobsters and oysters last night’. It would be very hard to argue that (3) is not a proposition, though it is made up of two propositions, it still is a proposition as much as either of its constituents. A similar trick can be done with our ‘or’ connective and the case for negation ‘ $\sim$ ’ is even easier to make as it is meant just to flip the truth value from true to false or false to true as the case may be. However when we arrive at the conditional things do not seem to pan out so easily.

Is what results from conditionalizing  $\psi$  on  $\theta$  a proposition? How we rule on this is pivotal to the debate regarding whether the conditionals that results has a truth values or not. If the  $\supset$  is more like  $\vee$  and  $\wedge$ , then the conditional may be truth functional because it is a proposition, if on the other hand the  $\supset$  is more like the modal operators  $\square$  and  $\diamond$ , it follows that

the conditional sentences they are found in, may not have truth values because they are not propositions. By stipulation ‘if  $\theta$  then  $\psi$ ’ is true, if it is not the case that  $\theta$  is true and  $\psi$  is false. Thus it follows, that the conditional is logically equivalent to  $\sim(\theta \rightarrow \sim\psi)$  and  $\sim\theta \vee \psi$ . The question then, is whether this is a good enough interpretation or rendering of ‘if  $\theta$ ,  $\psi$ ’. A conditional will be true (according to the truth functional account) when the components are, (true, true), (false, true) and (false, false). Thus any sentence of the form  $(\theta \rightarrow \psi) \rightarrow \theta$  is always true. This is obvious, if Alan and Bob have steak, then Alan has steak. The reason any sentence of this form is always true, is because the components are such that it would be impossible for it to have a true antecedent and a false consequent.

The truth functional account hinges on whether  $\theta \vee \psi$  entails ‘if  $\sim\theta$ ,  $\psi$ ’ or whether  $\sim(\theta \wedge \psi)$  entails ‘if  $\theta$ ,  $\sim\psi$ ’. Edgington argues that they do, she states that knowing “just that at least one of the propositions, ‘ $A$ ,  $B$ ’ is true is enough to infer that if  $\theta$  is not true,  $\psi$  is true” and further “knowing just that  $A$  and  $B$  are not both true is enough to infer that if  $A$  is true  $B$  is not.”<sup>8</sup> A strange consequence of this seems to be, that if we hold  $\sim\theta$  to be true, then it follows that a conditional of the form  $\theta \rightarrow \psi$  is also true. As bizarre as this may seem, it is hard to doubt.

If we find ‘he walked the dog’ to be false, then it seems to follow from that ‘if he walked the dog then he collected the mail’ is true. This may, as Edgington indicates, seem baffling but in strictly logical terms, it is permissible. This just might be a consequence of conditionalizing on false antecedents that we may need to accept. The conditional formed in this way may in fact be true but rather pointless, for practical purposes. Further bizarreness abounds when we consider the following sentences.

(1) 'I think my roommate isn't home yet, but if she is she'll be worried about where I am'

(2) 'I think the Prime Minister isn't home yet, but if he is, he'll be worried about where I am'

Now, (1) clearly is a reasonable thought to have, however (2) is perhaps a sign of madness for anyone not personally familiar with the Prime Minister. But why should it be? Indeed they have the same logical form and on the truth functional account they will have the same truth conditions. These examples due to Edgington, suggest that the truth functional account may have serious problems.

H.P Grice defended the truth functional account by drawing attention to statements which are false and those which are true but misleading. As most adults know, there are many ways of making a statement which is strictly speaking true however, somewhat misleading. Polite and civilized conversation is full of such statements. One easy way of doing this, is to state something weaker than that which you might be able to say. Edgington uses the following example. If we are asked 'where is Arthur?' We know for a fact he is in the casino and that he never goes to a church. We might then answer our interlocutor (not wanting to be rude) 'either he is in the casino or at the church'. Even though we know he is in the casino our disjunctive statement is still true. Equally we could have used a conditional 'if he's not in the bar then he is in the library'. As Edgington indicates, on the truth functional account I simply cannot doubt the truth of  $\theta \vee \psi$  if I already hold  $\theta$ .

Edgington finds this troubling and states that "we need to be able to discriminate believable from unbelievable conditions whose antecedent we think false. The truth functional account does not allow us to do this."<sup>9</sup> For this alone it may be unworthy of our allegiance. P.F Strawson held that if Grice is right about conditionals then the same can be said for the counterfactuals. Indeed, future oriented indicatives seem to behave like past oriented



counterfactuals (at least sometimes). Counterfactuals, with false antecedent are true, Grice explains, however, why some of them can be reasonable things to say while others cannot, this, according to him, can be accounted for through principles of good conversation. By contrast others such as Strawson have held that perhaps only in strict logical contexts can we hope to keep the meaning of  $\supset$  pure. Authors like James Thomson have agreed with Strawson's assertion finding that "it is possible to hold that in logic we adopt this interpretation of conditionals merely for convenience, because of certain given purposes so that the view is a kind of crude approximation of the facts."<sup>10</sup> What this means for the conditional and its study in natural language contexts isn't clear..

Other authors beginning with Jackson have attempted to secure the conditional by means of a notion of robustness. For Jackson, a conditional is robust when one would not abandon the conditional if one were to discover that the antecedent was true. Edgington puts it as follows "this condition is not satisfied if you believe  $A \rightarrow B$  solely on the grounds that  $\sim A$ . If you discovered that  $A$ , you would abandon your belief that  $A \rightarrow B$  rather than conclude that  $B$ ."<sup>11</sup> This directive comes from a simple intuition that most people have; at least in some cases, we seem to want to say it is the case that  $\sim \theta$ , but had it been the case that  $\theta$ , then  $\psi$  would have held as a consequence of  $\theta$ . We can see this illustrated quite easily in the following example, 'I am not a General in the US Air Force but had I been one, I would have been a member of the armed forces'. This seems to follow quite naturally however we should indeed be cautious with this type of an inference as it does not always produce such convivial results.

With regards to whether conditionals are propositions and if they are whether they can have truth values we seem to go in circles in some senses. We want to say that a conditional is a proposition constructed of two other propositions, and just like the connectives for 'and' and 'or'

a truth functional account of when the constituent parts are true, gives us an account of when the construct is true. Clearly, if we cannot resolve the question of embedded conditionals in larger constructions our account has failed. However perhaps we have gone astray in thinking that a conditional is a proposition to begin with, that is to say, that it is a sentence that sets up conditions that either succeed or fail and can in fact hold a truth value. Edgington puts it as follows, “we do not have a satisfactory account of sentences with conditional constituents. This may be because we have not yet figured out any truth conditions for conditionals. Or it may be because they don’t have any.”<sup>12</sup> We now turn to the related issue of whether or not conditionals can be seen as being part of fact stating discourse.

### **Conditionals and Fact Stating Discourse**

We normally like to think that given that there are propositions, that these propositions can represent the world. That is to say, we tend to think that a proposition like ‘Everest is the tallest mountain in the world’ does in fact represent a feature of the world. In the case of our sentence we believe it is true and that it reports back certain facts in the world to us. This is of course the nature of fact stating discourse and is tremendously important to anyone who takes the philosophy of science seriously. The question arises then, are conditionals part of this fact stating discourse that other bits of language seem to be part of, and indeed can they be if they are not propositions? All of this remains to be seen. Edgington suggests “that the mistake philosophers have made in trying to understand conditionals, is, to, treat them as part of fact stating discourse as representing the world as being a certain way and that this is not their

function.”<sup>13</sup> This suggests that perhaps our account of sentences with conditional constituents is unsatisfactory. This may be because we lack a proper account their truth conditions. Another possibility is that conditionals just don't have truth conditions. To the first possibility, perhaps conditionals are indeed proposition but the problem is in fact that for the natural language conditional, the truth conditions of Material Implication are somehow inappropriate. Indeed some have suggested that all we need to do to rehabilitate the natural language conditional is to reevaluate some of the aspects of the truth functional account. Most radically it has been suggested that for the truth table of the conditional of natural language, the last two lines ( 1)false antecedent true consequent gives true value for the conditional, and 2) false antecedent and false consequent gives true value for the conditional) should be dropped all together for natural language as they seem not to apply. To the second possibility, that conditionals do not have truth conditions because they are not propositions. Does this mean they cannot be part of fact stating discourse? Perhaps, but perhaps not, maybe the conditional can be part of fact stating discourse projectively, that is, as an epistemic strategy being projected on the world, we will have more to say about this shortly

The truth functional account at times seems at odds with common sense, the following argument according to the truth functional account of 'if' will be valid. 'If God does not exist, then it's not the case that if I pray my prayers will be answered. I do not pray therefore God exist.' We can show that this is valid quite easily in the following way. The argument has but one premise and we can represent it like this  $(\sim \theta \rightarrow \sim (\psi \rightarrow \chi)) \wedge \sim \psi$ , the conclusion is represented by the atomic expression  $\theta$ . Using elementary propositional logic we can show that the conclusion follows from the premise.

$$(\sim\theta \rightarrow \sim(\psi \rightarrow \chi)) \wedge \sim\psi$$

$$\therefore \theta$$

- |  |                                    |
|--|------------------------------------|
| (1) $(\sim\theta \rightarrow \sim(\psi \rightarrow \chi)) \wedge \sim\psi$ | [Assumption, premise]              |
| (2) $\sim\theta$   | [Assumption, Reductio ad absurdum] |
| (3) $(\sim\theta \rightarrow \sim(\psi \rightarrow \chi))$                 | [Simplification, 1]                |
| (4) $\sim(\psi \rightarrow \chi)$  | [Modus Ponens, 2,3]                |
| (5) $\sim\psi$   | [Simplification, 1]                |
| (6) $\sim\psi \vee \chi$   | [Or Introduction, 5]               |
| (7) $\psi \rightarrow \chi$  | [Material Implication, 6]          |
| (8) $\sim(\psi \rightarrow \chi) \wedge (\psi \rightarrow \chi)$           | [And Introduction, 4,7]            |
| (9) $\theta$   | [Reductio ad absurdum, 2,8]        |

QED

It is possible that my symbolic representation of the sentences ‘If God does not exist, then it’s not the case that if I pray my prayers will be answered. I do not pray therefore God exist’ as ‘ $(\theta \rightarrow \sim(\psi \rightarrow \chi)) \wedge \sim\psi$  therefore  $\psi$ ’ is not quite right. One might object to making ‘I do not pray’ ( $\sim\psi$ ) connected by an ‘and’ at the end and that it should not be one single premise but two premises ( premise 1: ‘ $(\theta \rightarrow \sim(\psi \rightarrow \chi))$ ’ and premise two: ‘ $\sim\psi$  ’ therefore: ‘ $\theta$  ’ ). This doesn’t seem to change the overall structure of the proof presented except by adding a line and some changes in the right hand side notation. The argument remains valid.

Others, have rejected this argument as inadmissible as it treats the existence of God in its subject matter. Very well, consider the following ‘If the tooth fairy does not exist, then it’s not the case that if I put my tooth under my pillow I will receive a quarter. I do not put my tooth

under my pillow therefore the tooth fairy exist'. Now clearly this is no longer a problem of whether God exists or not, but rather (and more seriously I would suggest) a problem of language and what valid arguments permit us to conclude. The problem isn't the content of the argument but the embedded conditionals in the antecedent. This shows that conditionals in the antecedent of other conditionals pose grave problems. Another example of this can be seen here 'if Marx was there if Engels was, then Bakunin was there'. As matter of course we have difficulty interpreting this symbolically. Should we interpret it like this ' $M \rightarrow (E \rightarrow B)$ ' or is it better interpreted by this ' $(M \wedge E) \rightarrow B$ '? A further question regards whether these two forms are equivalent. Edgington argues that "conditionals do not go into truth functional contexts or into each other easily" and what is more "those we do understand, e.g. conditionals in consequents, we understand as equivalent to sentences without embedded conditionals. The facts square at least as well with the hypothesis that conditionals do not have truth values as with the hypothesis that they do."<sup>14</sup> Some philosopher have argued that the fact that conditionals appear in valid arguments shows that conditionals must have truth conditions, otherwise, the valid arguments themselves would not be truth preserving. This conception of validity may in fact be drawn too narrowly for conditionals.

We may in fact want to re evaluate our account of the truth conditions of conditionals. This may be wiser than the full scale abandonment of the truth functional account of the conditional. The reason for this of course is related to the truth value itself, it seems hard to say that the conditional has a truth value if it isn't truth functional, however if a conditional already isn't a proposition then this shouldn't surprise or bother us tremendously. It is clear, that 'if  $\theta$ ,  $\psi$ ' is an assertion of  $\psi$  when  $\theta$  is true, and it asserts nothing at all when  $\theta$  is false. We naturally want to say that the conditional assertion is true if  $\theta$  and  $\psi$  are both true; and false, if  $\theta$  is true and

$\psi$  is false. Perhaps then it follows that the conditional has no truth value anytime that the antecedent is false.

Generally, most individuals understand that for consistency's sake, they cannot simultaneously believe both 'if  $\theta$ ,  $\psi$ ' and 'if  $\theta$ ,  $\sim \psi$ '. To accept 'if  $\theta$ ,  $\psi$ ' is to be inclined to reject 'if  $\theta$ ,  $\sim \psi$ '. Jackson has further argued that in understanding the nature of 'if' we may want to hold that the relationship between 'if  $\theta$ ,  $\psi$ ' and ' $\theta \rightarrow \psi$ ' can be said to be similar to the relation between ' $\theta$  and  $\psi$ ' and ' $\theta$  but  $\psi$ '. In this case we see 'but' as very much like 'and' however at the same time importing something further, a meaning related to contrast and comparison of two things. As in 'it's a good car but it's expensive' or 'he's smart but lazy'. Likewise with the 'if' of natural language, perhaps it is similar to  $\supset$  but with something more added to it. According to Edgington a "main source of our dissatisfaction with the truth functional conditional is a clash between our intuitions about validity and arguments it licenses as valid"<sup>15</sup> When we accept a conditional it is not the case that we express a belief in something rather it is that we hold a disposition to infer something from something else. D.H Mellor indicates that this may be appropriate, given that " 'If  $P, Q$ ' (i.e., one in which neither ' $P$ ' nor ' $Q$ ' contain conditionals) expresses a disposition to infer  $Q$  from  $P$ . In other words, fully to accept a simple 'If  $P, Q$ ' is to be disposed fully to believe  $Q$  if I fully believe  $P$ ."<sup>16</sup> Clearly, we shouldn't accept a conditional when its truth functional account is false but that doesn't mean that we should accept it uncritically when its truth functional account is true, many conditionals might be highly dubious, even if the material conditional which underpins them, is found to be true.

There have been many attempts to seek illumination to this question of fact stating discourse by appeal to possible world semantics. The results have varied in their success. Stalnaker developed in his logic a particular selection function which was designed to select

from any world, of the propositions in that world, the ‘nearest’ world where that proposition was true. This was done as way of projecting our epistemic strategies onto the world under consideration. Stalnaker argues for a “*selection function,  $f$* , which takes a proposition and a possible world as its value. The *s*-function selects for each antecedent *A*, a particular possible world in which *A* is true. The *assertion* which the conditional makes, then, is that the consequent is true in the world selected. A conditional is true in the actual world when its consequent is true in the selected world.”<sup>17</sup> We will have much more to say about possible worlds in a later section. Whether or not this selection function can help to determine the truth or falsehood of conditional statement remains an open question. If conditionals are indeed propositions then perhaps it can. Some theorists have argued that they are not, thus according to Edgington “there is no a priori reason why there should be a general routine for decoding compounds of them.”<sup>18</sup> Clearly, counterfactuals themselves cannot be part of fact stating discourse as they are ‘counter-to-the-facts’ that is, they are hypothesis contrary to fact. Whether an indicative conditional like ‘If the temperature drops low enough for long enough, then the pond in back of my house freezes’ can be part of fact stating discourse remains to be seen

When we consider counterfactuals and conditionals and whether they are part of fact stating discourse, it is perhaps helpful to lay out some grammatically different though factually (seemingly) equivalent sentences and judge how we are to evaluate them. We consider these three:

(BDD) *If John Wilkes Booth didn’t kill Abraham Lincoln, no one else did.*

(BDW) *If John Wilkes Booth doesn’t kill Abraham Lincoln, no one else will.*

(BHW) *If John Wilkes Booth hadn’t killed Abraham Lincoln, no one else would have.*

These are of course the famous ‘OK’ cases (Oswald/Kennedy) which have proven to be very useful as a sort of battle ground for many of the issues concerning conditionals and counterfactuals. They are due to Ernest Adams (1970). We will refer to these as DD, DW and HW respectively. Upon preliminary examination DD seems to be a past oriented indicative, DW seems to be a future oriented indicative and HW seems to be a past oriented counterfactual. Some (most notably Dudman) have been convinced that the difference between DW and HW is simply one of tense. For these theorists, HW just expresses what DW express, but at a later time. This is seen similarly in statements like ‘Tom is in London’ (expressed today) and ‘Tom was in London’ (expressed a week from now) This seems correct especially if we hold (as some have held) that these sentences express common place notions at play in grammar, Edgington suggests that we might view these sentences in the following way “ ‘would have’ is the past tense of will’ likewise for “ ‘should have/should’, ‘might have/may’” and “ ‘could have/can’ ”.<sup>19</sup> We can see then, that these tenses serve the function of clarifying temporal relations between the subordinate and main clauses of the conditionals in questions. We can all clearly understand that the if-clause must treat that which comes before or earlier (at the very least not later) than the main clause. Accordingly it is possible to explain why we seem to have a sense of tense shifting in the antecedent of DW to clarify the temporal relations, HW is of course the past tense of DW and has its antecedent in the past-past tense for the same reasons. As Sabine Iaridou has noted “*counterfactuality* is used as a term only with respect to situations that cannot be helped anymore.”<sup>20</sup> Edgington argues that by holding to this line of reasoning, in saying that HW is the past tense of DW we are committing ourselves to Bennett’s slogan concerning, them, that is, “every hadn’t-would was once a doesn’t-will.”<sup>21</sup> Edgington (though in agreement with Bennett) cautions us however not to draw the conclusion that to accept the ‘hadn’t-would’ conditional



means that earlier, we accepted the same ‘doesn’t-will’ conditional. Concerning Bennett’s account, Edgington for the most part agrees and finds that “for a large number of central cases the most commonly met and the most important counterfactuals” this will be true, however “for some counterfactuals it will be hard to find a temporal dimension.”<sup>22</sup> But, what of DD, there may be grounds for finding it to be a different type of conditional all together. Dudman held that DD should rightly be construed as some kind of condensed argument and that DW and HW as judgments or verdicts on the main clause which are being qualified by the subordinate clause.

There is, however reason to doubt this, for Edgington “to what extent you are actually doing a bit of inferring, as opposed to ‘imaginatively projecting’ (which Dudman sees as the process behind the verdict in DW and HW) – doesn’t necessarily coincide with the grammatical distinction.”<sup>23</sup> Classically philosophers have tried to keep the DDs separated from the DWs and the HWs and this has led to a shrinking of the class of indicatives and an expanding of the class of subjunctives. It has been assumed that the OK examples show us (because you can accept an HW while rejecting a DD) that there is really two distinct types of conditional and one is for talking about facts and the other is for talking about non facts. To this Edgington disagrees, because for her, all conditionals can be evaluated in terms of Ramseyian probability structures. She finds that if there is a distinction to be made between the three OK cases then it is not between DD and DW, which should be treated as a unified class, but between those two and HW. Still though, on her account the distinction isn’t really that significant anyway and as Dudman held, it is likely merely one of tense. For her then “DD and DW represent your present beliefs about how the world (actually) will be or was conditionally upon a supposition” where as “HW represents your belief about what was going to happen conditionally upon something else’s having happened which often you think actually didn’t.”<sup>24</sup> For Edgington then HW has a use

when we want to adopt an epistemic perspective different from our actual one. We take it to be suited to times in the past when we considered the antecedent as a genuine possibility.

Where does all of this leave us with regards to fact stating discourse? It's not clear. Of the three questions posed so far in this section, 1) The subjunctive indicative distinction, 2) Are conditionals propositions? 3) Are conditionals part of fact stating discourse? It really is the third one that has the least satisfactory answer either way, that is to say, either way this question is answered (they either are or aren't) it leaves a bad taste in ones mouth. Some have argued that the whole discussion of whether or not conditionals are part of fact stating discourse is rather pointless if you already hold that conditionals are not propositions. They contend that nothing that is not a proposition can be part of fact stating discourse. For my part I think this is a little too hasty. Just because conditionals may not be propositions doesn't mean they can't be part of fact stating discourse. It may be that we have drawn the bounds of fact stating discourse too narrowly by saying only propositions are part of fact stating discourse. I would argue, that conditionals are part of fact stating discourse, but that we use them projectively in the sense of projecting them on to the world as epistemic strategies for dealing with phenomenon. We would certainly be in a bad spot if it turned out a conditional like 'if you stop eating bacon and eggs for breakfast everyday you will improve the condition of your heart' wasn't at least in some sense factual, that is, representing real relations at play in the world between our diets and the conditions of our hearts. We not only 'like' conditionals like that, we indeed 'need' them. Until we somehow develop the ability to know all information about everything at all times (just like god I suppose) the conditional will remain part of our thought process. Whether or not it is just that, 'part of our thought process' or a fundamental feature that actually exists in the world independent of us, is a harder question to answer.

We have now completed our evaluation of the issues surrounding conditionals and counterfactuals and their inter relation. If it seems as though little has been resolved in terms of the issues raised then that sentiment is well founded. As previously indicated little consensus has emerged with regards to the most intractable problems presented here. We studied the indicative and subjunctive distinction and saw that despite the existence of some good reasons for collapsing the distinction our endeavors are best served by at least provisionally maintaining it. We also evaluated the question of whether conditionals are propositions and whether we should take them to fall under some kind of truth functional account and saw that it is problematic to consider them propositions and thus the ascription of truth values via a truth functional account was also problematic. Finally we evaluated in what respects conditionals can be seen to be part of fact stating discourse, here we saw the deepest disagreements with little or no consensus emerging. We will now move to evaluate the technical aspects of Lewis's system of counterfactuals.

## Chapter 2: Lewis's System.

David Lewis in his seminal 1973 book *Counterfactuals*, develops a highly sophisticated and detailed account of a logical system for counterfactual conditionals. The system is both complex and subtle and has become the center piece in the study of counterfactuals since its publication and has resulted in fierce debate with regards to its success and appropriateness in dealing with the pertinent issues raised by the study of these counterfactual conditionals. This chapter will attempt to recapitulate and evaluate Lewis' system as presented in *Counterfactuals*, focusing primarily on the first section *An Analysis of Counterfactuals*. In this effort we will aspire to remain as faithful as possible to the intentions of Lewis in laying down his system while reserving criticism for a later time.

### Lewis's System of Counterfactuals.

Lewis outlines the broad contours of his system in the first paragraph:

*'If Kangaroos had no tails they would topple over'* seems to me to mean something like this: in any possible state of affairs in which kangaroos have no tails, and which resembles our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over. I shall give a general analysis of counterfactual conditionals along these lines.<sup>25</sup>

To this end Lewis introduces two counterfactual conditional operators. Lewis attempts to interpret these operators by saying how the truth values at a given possible world where a counterfactual conditional of interest is applied will depend on the truth values at various possible worlds of both its antecedent and consequent terms. The structure of Lewis' system is founded on the fixed (though vague) parameter of 'comparative similarity', we shall return to this notion shortly in more detail. The two counterfactual conditional operators are that of the 'would' operator  $>\square$  and that of the 'might' operator  $>\diamond$ . Roughly speaking, they correspond to the two following sentence frame:

1)  $>\square$  *if it were the case that \_\_\_\_\_ then it would be the case that \_\_\_\_\_.*

2)  $>\diamond$  *if it were the case that \_\_\_\_\_ then it might be the case that \_\_\_\_\_.*

Both counterfactual conditional operators can be conditionalized on atoms or full sentences, they can also be embedded. This can be seen in the following well formed formulae:

$$\chi >\square \psi$$

$$(\chi >\square \psi) >\square \psi$$

$$(\chi >\square \theta) >\diamond (\chi >\square \psi)$$

$$((\psi >\square ((\chi >\square \psi) >\diamond \theta)) >\diamond \chi) >\square (\theta >\square (\psi >\diamond ((\chi >\square \theta) >\diamond (\theta >\square \psi))))$$

The two counterfactual operators can be seen to be to be interdefinable in the following way:

$$\theta >\diamond \psi =^{\text{df}} \sim (\theta >\square \sim \psi)$$

$$\theta >\square \psi =^{\text{df}} \sim (\theta >\diamond \sim \psi)$$

Lewis takes the ‘would’ counterfactuals ‘ $\>\square$ ’ as primitive to his system. With regards to the term ‘were’, at least in English this term has both a subjunctive and a temporal element to its meaning. There is of course also in English a tendency to view the antecedent as being invariably false. Lewis argues that in light of this we can see that “the counterfactual construction of English do carry some sort of presupposition that the antecedent is false” however “it is some sort of a mistake to use them unless the speaker does take the antecedent to be false and some sort of mishap to use them when the speaker wrongly takes the antecedent to be false.”<sup>26</sup> This may merely be an idiosyncratic feature of English and may not be anything deeper or more remarkable than that. There is hardly any reason to suppose that every type of presupposition failure must inevitably produce an automatic falsity or any kind of truth value gap for counterfactuals. Lewis suggests that this may be a consequence of conversational implicature of everyday speech with no effect on the truth conditions as we saw earlier this is somewhat similar to what Grice had in mind. Authors like Kit Fine have argued that this commits Lewis to the view that an inference like the following must be sanctioned:

$\theta, \psi$

---

$\therefore \theta \>\square \psi$

According to Fine, if “ $\theta$  and  $\psi$  are both true then the actual world is a  $\theta$ -world and not  $\theta \& \sim \psi$  world is as close, since no world is as close to a given world as that world itself”<sup>27</sup> Lewis accounts for this by holding that one can indeed assert a counterfactual whose antecedent is known to be true this will lead to a statement that is clearly misleading but not obviously false. Fine suggests that at least sometimes it may be appropriate to assert a counterfactual while in ignorance of the antecedent’s truth value. He presents the following as an example to consider, “I may speculate on a student’s prospects in an exam the result of which are already settled and

assert: ‘If he had worked hard he would have passed.’ My assertion is false if the student worked hard but was only able to pass through cheating.”<sup>28</sup>

The kind of conditional the counterfactuals represent is, and how we rule on this, is deeply important to the development of the system. Lewis holds that the counterfactual conditional cannot be seen to be any type of strict conditional. A strict conditional is a material conditional preceded by a necessity operator ( $\Box$ ).

$$\Box(\theta \supset \psi)$$

Within Lewis’ system we see that the necessity operator and the possibility operator are interdefinable in ways that should be familiar:

$$\Diamond\theta =_{df} \sim \Box \sim \theta$$

and

$$\Box\theta =_{df} \sim \Diamond \sim \theta$$

This interdefinability is what is known as the dual form. The strict conditional can also be rewritten using the possibility operator.

$$\sim \Diamond (\theta \wedge \sim \psi)$$

The necessity operator is in general seen to be like a restricted universal quantifier over possible worlds, it ranges over possible worlds that satisfy certain conditions and restrictions. The ‘worlds’ are accessible, which means they satisfy the restrictions and conditions under consideration. Necessity is truth at all the accessible worlds and different sorts of necessity correspond to different accessibility restrictions. The same can be said for the possibility

operator, possibility is truth at some accessible world. Accessibility can be defined in the following way. Typically accessibility is seen as a relative matter, it is a relationship between worlds. If we use the template of physical necessity, then by allowing  $i$  and  $j$  be worlds with different laws of nature, further we allow for a world  $k$  to be a world where we find the laws of  $i$  to hold true where laws of  $j$  are to be violated. From the point of view of  $i$ ,  $k$  is thus accessible but from the point of view of  $j$  it is not. We can thus say that  $k$  is accessible from  $i$ , however  $k$  is not accessible from  $j$ . Accessibility serves to restrict quantification over worlds by giving truth conditions that are either satisfied or not. The truth conditions for  $\Box\theta$  and  $\Diamond\theta$  are given by the following stipulation. For any world  $i$  and sentence,  $\theta$  the sentence  $\Box\theta$  is true at world  $i$ , if and only if, for every world  $j$  such that  $j$  is accessible from  $i$ ,  $\theta$  is true at  $j$ . Similarly  $\theta$  is true at  $i$  if and only if, for some world  $j$  such that  $j$  is accessible from  $i$ ,  $\theta$  is true at  $j$ . In other words  $\Box\theta$  is true at  $i$ , if and only if  $\theta$  is true at every world accessible from  $i$ ,  $\Diamond\theta$  is true at  $i$  if and only if  $\theta$  is true at some world accessible from  $i$ . Lewis chooses to amend this classical ascription of accessibility preferring rather an equivalent formulation based on what he calls *spheres*. Simply put, the spheres form a system such that an assignment to each world  $i$ , of a set  $S_i$  of worlds, which is called the *sphere of accessibility* around world  $i$  and should be regarded as the set of worlds accessible from  $i$ .

Given this system of spheres, we can give an amended account of the truth conditions of the modal sentences. A modal sentence  $\Box\theta$  is true at a world  $i$  if and only if  $\theta$  is true throughout a sphere of accessibility  $S_i$  around  $i$ . A modal sentence  $\Diamond\theta$  is true at a world  $i$  if and only if  $\theta$  is true somewhere in the sphere  $S_i$  around  $i$ . A strict conditional  $\Box(\theta \rightarrow \psi)$  is true at  $i$  if and only if  $(\theta \rightarrow \psi)$  is true throughout the sphere  $S_i$ , that is to say if  $\psi$  is true at every  $\theta$ -world in  $S_i$ . Different types of accessibility assignments for various types of necessity might be used



depending on the interest we have in formulating the system. The following are some of the possible accessibility relations that might be used:

- (1) Logical Necessity
- (2) Physical Necessity
- (3) Inevitability at time  $t$
- (4) Necessity in respect to facts of so and so kind
- (5) Necessity in respect to all facts
- (6) Deontic necessity
- (7) Vacuous necessity

### **Strict Conditionals and Variably Strict Conditionals**

With regards to the strictness of conditionals, the more inclusive the spheres of accessibility they operate in, the stricter the conditional must be. Of course it follows that any strict conditional is implied by a stricter conditional with the same antecedent and consequent. Lewis describes the corresponding strictness of various conditionals and their interrelations, “the logical strict conditional is stricter than any other; the material conditional is the least strict of all the conditionals that obey the constraint that every world is self-accessible; and the physical strict conditional, for instance falls in between. The vacuous conditional is the least strict conditional of all.”<sup>29</sup> For their part, counterfactuals are related to a type of strict conditional based on the comparative similarity of possible worlds.

Counterfactuals can be related to a kind of strict conditional by means of a notion of comparative similarity of a possible world. A counterfactual  $\theta >_{\square} \psi$  is true at a world  $i$  if and only if  $\psi$  holds at certain  $\theta$ -permitting worlds, clearly, not all the  $\theta$ -permitting worlds matter. Lewis puts it as follows “‘*if kangaroos had no tails, they would topple over*’ is true (or false, as the case may be) at our world, quite without regard to those possible worlds where kangaroos walk around on crutches and stay upright that way. Those worlds are too far away from ours.”<sup>30</sup>

What we mean by the counterfactual is that things being more or less as they are if kangaroos had no tails they would indeed fall over. Our intuitions incline us towards directing our attention to worlds in which the kangaroos have no tails but everything else remains the same. However, this is somewhat deceptive as there are no such worlds given that if there were a world where kangaroos had no tails, certain other facts would also have to be revised as well, for example if they had no tails their tracks would likewise would be different and similarly their genetic code would also be different. Perhaps then, we should confine our attention only to the more salient features that bear on the question at hand. As Lewis puts it, “respects of similarity and difference trade off. If we try too hard to exact similarity to the actual world in one respect, we will get excessive differences in some other respect.”<sup>31</sup> It seems then that counterfactuals are strict conditionals corresponding to our accessibility assignment determined by similarity of worlds, the overall similarity with respect to differences balanced off against respects of similarity. However the notion of comparative similarity has engendered vigorous debate, Kit Fine has argued that “the notion of comparative similarity gives rise to an immediate danger of circularity. For similarity is a matter of agreement in propositions; and among those propositions will be counterfactual ones. So to evaluate a counterfactual one needs to compare worlds for similarity to the actual world and this would seem to require the evaluation of further

counterfactuals.”<sup>32</sup> Furthermore, Fine argues that “what makes Lewis’s theory hard to judge is that the notion of similarity tends to accommodate itself to the counterfactual at hand; that the sense of similarity is chosen that will yield the correct truth conditions. The question is whether this accommodation can be characterized without circularity” Fine concludes that it is “not clear if similarity can be specified without circularity or without appeal to some independent theory of counterfactuals.”<sup>33</sup> Does this mean that Lewis’ account is flawed from the outset as it seems to be founded on a circular notion of similarity? Some have argued that it does, however perhaps we should bracket this criticism for the meantime while we pursue our evaluation of the technical aspects of Lewis’ system.

If we take the set  $S_i$  for each world  $i$ , that is similar within a prescribed and stipulated degree of similarity the corresponding strict conditional of any counterfactual is true if and only if the material conditional holds throughout the set, that is to say if and only if the consequent holds at all the antecedent worlds similar to that degree with  $i$ . This seems to work well when we consider only one counterfactual but it seems to break down when we consider more than one. The following example furnished by Lewis seems to show this quite well:

If the USA threw its weapons into the sea tomorrow there would be war; but if the USA and the other nuclear powers all threw their weapons into the sea tomorrow there would be peace; but if they did so without sufficient precautions against polluting the world’s fisheries there would be war; but if after doing so, they immediately offered generous reparations for the pollution there would be peace...<sup>34</sup>

We can represent these sentences symbolically in the following way



level. Lewis argues that “if counterfactuals are strict conditionals we have no hope of deciding once and for all how strict they are”<sup>37</sup> Furthermore, no context “can favor a strictness on which the four sentences from the pairs at two adjacent stages are all true” the reason for this as Lewis puts it, is because there “is no such strictness.”<sup>38</sup> There may be some grounds for holding that counterfactuals are a type of vague strict conditional based on similarity and what is more the vagueness can be resolved by means of fixing the strict conditional by local context given by the antecedent itself. However, Lewis argues that although this may not be completely wide of the mark it may however be, so to speak, to throw the baby out with the bath water, when a systemic analysis may still have some hopes of being created.

The answer may be as Lewis suggests, that there are grounds to hold that the counterfactual is not any particular strict conditional, but rather, might be what could be called, a ‘variably strict conditional’. Lewis suggests that “any particular counterfactual is as strict within limits, as it must be to escape vacuity, and no stricter.”<sup>39</sup> When we consider a strict conditional, we take an assignment to each world  $i$  of just one sphere of accessibility  $S_i$  around a world  $i$ . In contrast when we consider a variably strict conditional we must assign to each world of a set  $\$_i$  of the spheres of accessibility around  $i$ .

We can define such an assignment for the variably strict conditional in the following way. We allow for  $\$$  to be an assignment to each possible world  $i$  of a set  $\$_i$ , of sets of possible worlds. It follows that  $\$$  is a ‘centered system of spheres’, each of the members of each  $\$_i$  are to be called spheres around  $i$ , if and only if for each world  $i$ , the following conditions hold:

(C)  $\$_i$  is *centered on*  $i$ ; that is the set  $\{i\}$  having  $i$  as its only member belongs to  $\$_i$ .

(1)  $S_i$  is *nested*; that is whenever  $S$  and  $T$  belongs to  $S_i$ , either  $S$  is included in  $T$  or  $T$  is included in  $S$ .

(2)  $S_i$  is *closed under unions*; that is, whenever  $\delta$  is a subset of  $S_i$  and  $\bigcup \delta$  is the set of all worlds  $j$  such that  $j$  belongs to some member of  $\delta$ ,  $\bigcup \delta$  belongs to  $S_i$ .

(3)  $S_i$  is *closed under (nonempty) intersections*; that is whenever  $\delta$  is a nonempty subset of  $S_i$  and  $\bigcap \delta$  is the set of all worlds  $j$  such that  $j$  belongs to every member of  $\delta$ ,  $\bigcap \delta$  belongs to  $S_i$ .<sup>40</sup>

This system of spheres is used to interpret counterfactuals, such that, it can carry information about the comparative and overall similarity of worlds. It should be understood then, that the smaller the sphere the more similar to  $i$  a world must be to fall within in it. More succinctly we can say that if we have a world that is within a sphere around a world  $i$ , and another world outside of that sphere, the world within the sphere is more similar to  $i$  than the world outside of the sphere. Additionally we can say that “if  $S$  is any set of worlds such that every member of  $\delta$  is more similar to  $i$  than any non member of  $S$ , then  $S$  should be one of the spheres around  $i$ .”<sup>41</sup>

### **The System of Centered Spheres**

The four formal constraints in the definition laid out for a system of centered spheres can be seen to be justified because if they were not fulfilled the system of spheres could not be regarded as carrying information about the comparative similarity of worlds. We can say more about each condition individually. Firstly, for condition (C), it is obvious that any world  $i$  must be as similar to itself as any other world is to it; thus it follows that  $i$  should belong to every non

empty sphere around  $i$ . It also follows from this definition that no other world is as similar to a world  $i$ , as  $i$  itself. This applies even to worlds which are qualitatively indiscernible from  $i$ . There should be some sphere around  $i$  that contains  $i$  and no other world, that is to say it, will be the sphere of  $\{i\}$  around  $i$ . Secondly, for condition (1) if it were the case that some  $\$i$  were not nested, it would follow that we would have two spheres  $S$  and  $T$  within  $\$i$ , and two worlds  $j$  and  $k$ , such that  $j$  would lie inside the sphere  $S$  but outside the sphere  $T$ , and  $k$  then would lie within  $T$  but outside  $S$ . If  $S$  and  $T$  both carry information about the comparative similarity to  $i$ , then it would follow that  $j$  must be more similar than  $k$  to  $i$ . This is the case because  $j$  is within the sphere  $S$  and  $k$  is not. However  $k$  would be more similar than  $j$  to  $i$ , this is the case because  $k$  is within the sphere  $T$  and  $j$  is not. It seems that it cannot be both and thus must be one or the other. Thirdly, for condition (2) if we assume that  $j$  lies within the union  $\cup \delta$  of the set of  $\delta$  and  $k$  does not. It should follow then from this that  $j$  lies within some sphere  $S$  in  $\delta$  and  $k$  does not. Additionally then it follows that  $j$  is more similar than  $k$  to  $i$ . We can conclude then that  $\cup \delta$  is a set for which any world that is found within it is more similar to  $i$  than any other world which is outside it, and that set should be a sphere around  $i$ . Fourthly, for condition (3), like condition (2) if we assume  $j$  lies within the intersection  $\cap \delta$  of a nonempty set  $\delta$  of spheres, and  $k$  does not lie within the intersection  $\cap \delta$ , then it follows that  $j$  lies within some sphere  $S$  in  $\delta$  and  $k$  does not. Thus  $j$  is more similar than  $k$  to the world  $i$ .  $\cap \delta$  then is a set which any world inside it is more similar to  $i$  than any world which is outside it, and that set should be a sphere around  $i$ .

We can draw some broader conclusion about these conditions and their interrelationship. Conditions (2) and (3), which give us closure under unions and intersections will be automatically satisfied when there are only finitely many spheres around  $i$ , or also in the more restricted case of a finite subset  $\delta$  of an infinite  $\$i$ . Lewis asserts “if there is a biggest sphere in  $\delta$

(one that includes all the others) it is  $\bigcup \delta$ . If there is a smallest sphere in  $\delta$  (one that is included in all others) it is  $\bigcap \delta$ .<sup>42</sup> By means of nesting, a finite set has a biggest and a smallest sphere, however this is not so for infinite sets which will possibly have no biggest or smallest sphere. Questions seem to arise with regards to whether or not union  $\bigcup S_i$ , of all spheres around  $i$  should exhaust the set of worlds, that is to say whether or not every possible world is to lie within some sphere around  $i$ . Regarding this, Lewis argues “if  $\bigcup S_i$  is the set of all worlds for each  $i$ , I will call  $S$  universal. If not then I regard the worlds the spheres around  $i$  do not reach- those that lie outside  $\bigcup S_i$  - as being all equally similar to  $i$ , and less similar to  $i$  than any world that the spheres do reach. We will see that any such world will be left out of consideration in determining whether a counterfactual is true at  $i$ . It is as if, from the point of view of  $i$ , these remotest worlds were not possible worlds at all.”<sup>43</sup>

### **Truth Conditions for Counterfactuals**

Having established his system of spheres Lewis moves on to give precise truth conditions to his counterfactual conditionals as follows.

$\theta > \square \psi$  is true at world  $i$  (according to a system of spheres  $S$ ) if and only if, either:

- (1) no  $\theta$ -world belongs to any sphere  $S$  in  $S_i$ , or
- (2) some sphere  $S$  in  $S_i$  does contain at least one  $\theta$ -world, and  $\theta \rightarrow \psi$  holds at every world in  $S$



Condition (1) gives the vacuous case, in this case either we will find that  $\theta$  is true at no world, or it is true only at worlds found outside of  $\cup \mathcal{S}_i$ . In this case the counterfactual is vacuously true at  $i$ . We can say that the statement  $\theta$  is not entertainable at  $i$ , as a counterfactuals supposition. Condition (2) gives the principal case, in this case,  $\theta$  is an entertainable supposition at  $i$ , and inside some sphere around  $i$ , which is large enough to reach at least one  $\theta$ -world. We can call such a sphere a  $\theta$ -permitting sphere,  $\psi$  will be thus true at all such  $\theta$ -worlds. Lewis puts it as follows, in brief; “a counterfactual is vacuously true if there is no antecedent-permitting sphere, non vacuously true if there is some antecedent-permitting sphere in which the consequent holds at every antecedent-world and false otherwise.”<sup>44</sup>

From these two conditions four cases emerge for a counterfactual  $\theta >_{\square} \psi$ . Two of these are cases where the counterfactual is to be true at  $i$ , and two are ways for it to be false. We will take the time to examine each of these cases in turn.

Case (A) is such that there is no  $\theta$ -permitting sphere. In this case we have vacuous truth. Every counterfactual with the antecedent  $\theta$  will be vacuously true at world  $i$ , this is so because the outermost of the spheres around world  $i$  does not reach the  $\theta$  world, indeed there may not be any such world. In this case both  $\theta >_{\square} \psi$  and  $\theta >_{\square} \sim \psi$  will be found to be true.

Case (B) is such that there is a  $\theta$ -permitting sphere around  $i$  within which  $\psi$  holds at all  $\theta$ -worlds, namely the next to outermost sphere  $\theta \rightarrow \psi$  holds throughout this sphere. In this case we have non vacuous truth. It should be noted that only one such sphere is necessary to make the counterfactual true. In this case both  $\theta >_{\square} \psi$  and  $\sim (\theta >_{\square} \sim \psi)$  will be found to be true.

Case (C) is such that there are  $\theta$ -permitting spheres but none in which  $\psi$  holds at every  $\theta$ -world, but no spheres throughout which  $\theta \rightarrow \psi$ . In this case we have falsity, opposite of true. In this case both  $\sim (\theta >_{\square} \psi)$  and  $\theta >_{\square} \sim \psi$  will be found to be true.

Case (D) is such that there are  $\theta$ -permitting spheres and both of them contain a mixture of  $\theta$ -worlds where  $\psi$  holds and  $\theta$ -worlds where  $\sim \psi$  holds. In this case we have falsity, opposite of false. In this case both  $\sim (\theta >_{\square} \psi)$  and  $\sim (\theta >_{\square} \sim \psi)$  will be found to be true.

Now we can reconsider our sequence of true counterfactuals and their negated opposites, which was the motivation for discarding the claim that the counterfactual is a constantly strict conditional founded on the notion of similarity. If we look again at our sequences of counterfactual statements:

$\theta_1 >_{\square} \psi$                       and     $\sim (\theta_1 >_{\square} \sim \psi)$

$\theta_1 \wedge \theta_2 >_{\square} \sim \psi$               and     $\sim (\theta_1 \wedge \theta_2 >_{\square} \psi)$

$\theta_1 \wedge \theta_2 \wedge \theta_3 >_{\square} \psi$               and     $\sim (\theta_1 \wedge \theta_2 \wedge \theta_3 >_{\square} \sim \psi)$

We remember the trouble of taking the counterfactual as a constantly strict conditional was that we had to choose one of the spheres  $S_i^1, S_i^2, S_i^3$  of each level to be the appropriate sphere of accessibility around world  $i$ . We saw that no choice was right, if we choose  $S_i^1$  for the first stage it was not right for the second,  $S_i^2$ , if we choose  $S_i^2$  for the second stage it was not right for the

third  $S_i^3$ , and so on at infinitum. The solution of course is to take the counterfactual as a variably strict one, if we do this then we have no need to choose. All of these respective spheres will be found in  $S_i$ . Each of the  $S_i^n$  will be found there in the set  $S_i$  to make the counterfactual non-vacuously true at each level. Each stage can now exist in harmony with the next one.

### **The Limit Assumption**

We will now proceed to examine one of the most important principles of Lewis' system that of the Limit Assumption. When we have only a finite number of spheres around a world  $i$ , then it follows that any nonempty set of spheres will have a sphere that is the smallest sphere of them all, which will be found in the set and will be included in all the other spheres in this set. Lewis argues that "for any entertainable antecedent the set of antecedent-permitting spheres has a smallest member. This smallest antecedent-permitting sphere is the intersection of the set of all antecedent -permitting spheres around  $i$ . It contains the antecedent worlds closest to  $i$ : all and only those of the antecedent-worlds than which no other antecedent world is closer to  $i$ ."<sup>45</sup> If we grant that there are sequences of ever smaller spheres without end, then it follows that we must have sets of sets of sphere with no smallest member. However it still might be the case that for every entertainable antecedent in the language, there will be a smallest antecedent permitting sphere. The Limit Assumption is the assumption that we can take smaller and smaller antecedent permitting spheres containing antecedent worlds closer and closer to  $i$ . However this process cannot go on forever for "we eventually reach a limit: the smallest antecedent permitting sphere and in it the closest antecedent-worlds" additionally "if the consequent of a counterfactual holds

at all antecedent worlds within some antecedent-permitting sphere around  $i$ , then also the consequent holds at all antecedent-permitting sphere.”<sup>46</sup> By the limit assumption we can simplify the truth conditions for counterfactuals as follows:

A counterfactual is true at  $i$  if and only if either

- (1) There is no antecedent-permitting sphere around  $i$ .
- (2) The consequent holds at every antecedent world in the smallest antecedent permitting sphere around  $i$ .

However, nothing warrants our assumptions that there always is a smallest antecedent permitting sphere and within it a set of closest antecedent permitting worlds. We are presented with a vicious infinite regress in some cases when we compare certain objects in our world to possible worlds. Lewis uses the example of line segments. If we have a line about an inch long in world then presumably if we have other worlds with a line that is 2 inches long and world with a line that is one and half inches long, then the second line segment is closer than the first to our world. However, a world with a line segment that is one and one quarter inches will be still closer. Of course this process can go on indefinitely in Zenoese fashion, it's clear that this phenomenon will apply to many other quantitative measures like area, volume and indeed lengths of time like minutes and hours. In these cases we may need to admit that there is in fact no such thing as a closest world. When there is no smallest antecedent permitting sphere the truth conditions for counterfactuals are as follows; “if there are antecedent-permitting spheres,

then as we take smaller and smaller ones without end, eventually we come to ones in which the consequent holds at every antecedent-world.”<sup>47</sup>

We will now move to examine two further features of Lewis’ system, that is, his interpretation of the ‘might’ counterfactuals and the outer modalities. Lewis gives an interpretation of the might counterfactual as a variably strict conditional. If we remember Lewis has defined the ‘might’ counterfactual in the following way:

$$\theta > \diamond \psi =_{df} \sim (\theta > \square \sim \psi)$$

As we see, the ‘might’ counterfactual is defined in terms of the ‘would’ counterfactual and they yield the following derived truth conditions.

$\theta > \diamond \psi$  is true at a world  $i$  (according to a system of spheres  $\$$ ) if and only if both

- (1) some  $\theta$ -world belongs to some spheres  $S$  in  $\$_i$  and
- (2) every sphere  $S$  in  $\$_i$  that contains at least one  $\theta$ -world contains at least one world where  $\theta \wedge \psi$  holds.

By the limit assumption we can re state this truth conditions of the ‘might’ counterfactual in the following way “ $\theta > \diamond \psi$  is true at  $i$  if and only if  $\psi$  holds at some  $\theta$ -world in the smallest  $\theta$ -permitting sphere around  $i$ .”<sup>48</sup> More simply put a ‘might’ counterfactual is true at  $i$  if only if the

consequent is true at world  $i$  if and only if the consequent holds at some antecedent-world closest to  $i$ . However if the limit assumption does not hold this restatement will prove unsatisfactory and the ‘might’ counterfactual is then found to be true if and only if as we take smaller and smaller antecedent permitting spheres around  $i$  without any end, and this will confine our attention to the antecedent-worlds closer and closer to  $i$ , we find however that we can never leave behind all the antecedent-worlds where the consequent is found to hold. If the ‘would’ counterfactual  $\theta > \square \psi$  is non-vacuously true then the ‘might’ counterfactual  $\theta > \diamond \psi$  will be found to also be true. If  $\theta > \square \psi$  and its corresponding opposite  $\theta > \square \sim \psi$  are both false, then it follows that  $\theta > \diamond \psi$  and its opposite  $\theta > \diamond \sim \psi$  are consequently both true. This is the case where we find that  $\psi$  is true at some of the closest  $\theta$ -worlds and  $\sim \psi$  is found to be true at others. However when  $\theta > \diamond \psi$  is false and its corresponding opposite  $\theta > \diamond \sim \psi$  is true  $\psi$  holds at none of the closest of the  $\theta$ -worlds and  $\theta > \diamond \psi$  is thus false. Additionally when  $\theta$  is not entertainable and  $\theta > \square \psi$  is thus vacuously true and  $\theta > \diamond \psi$  is thus false.

Suppose ‘ $\top$ ’ and ‘ $\perp$ ’ as sentential constants, true and false at every possible world respectively. We can define them in the following ways:

$\top$ : can be any arbitrary truth functional tautology

$\perp$ : can be its contradictory negation.

Then, in terms of the ‘would’ counterfactual we can say that  $\theta > \square \perp$  cannot be true except when vacuously true, that is when  $\theta$  is not entertainable

Then, in terms of the ‘might’ counterfactual, we can say that  $\theta > \diamond \top$  by definition is equivalent to  $\sim (\theta > \square \sim \top)$  and thus the negation of  $\theta > \square \perp$  is true if and only if  $\theta$  is entertainable.

### Modal Interpretation of the System

We can give the following modal interpretation, with the familiar modal operators which can be defined in terms of the counterfactual conditional connectives.

$$\diamond \theta =^{\text{df}} \theta > \diamond \top \text{ or } (\sim (\theta > \square \perp))$$

$$\square \theta =^{\text{df}} \sim \diamond \sim \theta \text{ or } (\sim \theta > \square \perp)$$

We can read  $\diamond$  and  $\square$  in following ways:

$\diamond$ : ‘it is entertainable that \_\_\_\_\_’

$\square$ : ‘it would be the case no matter what that \_\_\_\_\_’

That is to say possible and necessary in the usual parlance of modal logic.

Additionally the two modal operators are also inter definable in terms of the familiar dual form:

$$\diamond \theta =^{\text{df}} \sim \square \sim \theta$$

$$\square \theta =^{\text{df}} \sim \diamond \sim \theta$$

Finally we can give a definition of the modal operators in terms of the counterfactual connectives in the following way:

$$\diamond\theta =_{df} \theta > \diamond\theta$$

$$\square\theta =_{df} \sim\theta > \square\theta$$

These last definitions are equivalent to the dual form.

From these truth conditions for counterfactuals and the definition of the two modal operators we can further derive appropriate truth conditions for modal sentences. These are as follows:

$\diamond\theta$  is true at world  $i$  (according to a system of spheres  $\$$ ) if and only if  $\theta$  is true at some world in some sphere  $S$  in  $\$_i$ .

$\square\theta$  is true at world  $i$  (according to a system of spheres  $\$$ ) if and only if  $\theta$  is true at every world in every sphere in  $\$_i$ .

Given this, we can re express these truth conditions in terms of the assignment to each world  $i$  of the set of worlds  $\cup \$$ , (which is the union of all spheres around  $i$ )  $\cup \$_i$  is understood to be a sphere around  $i$ , indeed it is the largest or outermost sphere around  $i$ . We can express these truth conditions as follows:



$\diamond\theta$  is true at  $i$  (according to  $\mathcal{S}$ ) if and only if  $\theta$  is true at some world in  $\cup \mathcal{S}_i$ , this can be seen to be analogous to the existential quantifier.

$\Box\theta$  is true at  $i$  (according to  $\mathcal{S}$ ) if and only if  $\theta$  is true at every world in  $\cup \mathcal{S}_i$ , this can be seen to be analogous to the universal quantifier.

Given these truth conditions and definitions, according to Lewis “our defined modal operators turn out to be interpretable in the usual way by means of accessibility they correspond to each world  $i$  of the single sphere of accessibility  $\cup \mathcal{S}_i$ ” it follows then that “they pertain to the outermost of our spheres around each world  $i$ .”<sup>49</sup> We can call these the *outer modalities*, and these correspond to outer necessity and outer possibility. Either the system is universal or not, if it is then the necessity and possibility is that of the ordinary logical kind, if not, varying degrees of strictness are possible for the modalities.

We can see that our reading of  $\diamond\theta$  as ‘*it is entertainable that  $\theta$* ’ is justified because this is the case when there is some  $\theta$ -permitting sphere around  $i$  so that a counterfactual with  $\theta$  as its antecedent can be false or non vacuously true at  $i$ , that is, if and only if  $\theta$  is true at some world in some sphere around  $i$ . Likewise, our reading of  $\Box\theta$  as ‘it would be the case no matter what that  $\theta$ ’ this is the case, when  $\Box\theta$  is true at a world  $i$ , then  $\chi >_{\Box} \theta$  is true at  $i$  for any antecedent  $\chi$ . When  $\chi$  is not entertainable then  $\chi >_{\Box} \theta$  is vacuously true. If on the other hand  $\chi$  is entertainable, then  $\chi >_{\Box} \theta$  is non vacuously true because by hypothesis  $\theta$  is true throughout every sphere around  $i$  and thus is true throughout some  $\chi$ -permitting sphere around  $i$  and therefore  $\chi \rightarrow \theta$  is true

throughout some  $\chi$ -permitting sphere around  $i$ .  $\Box(\theta \rightarrow \psi)$  is the outer strict conditional it implies the counterfactual conditional  $\theta >\Box \psi$ . When  $\theta \rightarrow \psi$  is true throughout every sphere around  $i$ , then if there is any  $\theta$ -permitting sphere around  $i$ , it is true everywhere throughout that sphere. However it should be stressed that the converse does not hold.  $\theta >\Box \psi$  is true and  $\Box(\theta \rightarrow \psi)$  is false if  $\theta \rightarrow \psi$  is true throughout some  $\theta$ -permitting sphere, but it is false somewhere in some larger  $\theta$ -permitting sphere.

### Impossible Antecedents

We will now move to examine another aspect of Lewis' system, that of the impossible antecedent. There is some justification at the intuitive level for having a 'would' counterfactual with an impossible antecedent to come out to be vacuously true. A counterfactual with an antecedent which logically implies the consequent should always be true. On the other hand a counterfactual with an impossible antecedent, a self contradictory one, can imply any consequent. As Lewis puts it "along with the *possible* possible worlds that differ from our world only in matters of contingent empirical fact there also are some *impossible* possible worlds that differ from our world in matters of philosophical, mathematical and even logical truth."<sup>50</sup> Lewis' account makes all of these counterfactuals come out vacuously true. However it may not be necessary to discriminate truth values between possible and impossible counterfactual even if some may be assertable and some may not. If we examine the following sets of counterfactuals with impossible antecedents then we can see that some might be acceptable while others are dubious. The following show this quite well:

- 1) *If there were a largest prime  $p$ ,  $p! + 1$  would be prime.*
- 2) *If there were a largest prime  $p$ ,  $p! + 1$  would be composite.*

These two seem sensible. Whereas:

- 3) *If there were a largest prime  $p$ , there would be six regular solids.*
- 4) *If there were a largest prime  $p$ , pigs would have wings.*

Seem not.

We could easily extend our first list (1 and 2) of counterfactuals with impossible antecedents that we can easily accept further by adding sentences like:

- 5) *If 4 were not divisible by two, then 4 would be a prime number*
- 6) *If figure 'A' were a regular polygon then figure 'A' would be symmetrical*

This process of creating acceptable counterfactual statements with impossible antecedents isn't too hard, all we need to do is take a definition from mathematics and negate one of its components and put it in counterfactual conditional form. Likewise our second list (3, and 4) could be extended in the following way:

- 7) *If there were a smallest positive real number then Exxon Mobile would go out of business*

8) *If there were any other number besides '26' that was sandwiched between a perfect square (25) and a perfect cube (27), then Bob is your uncle.*

Again all we seem to be doing is taking a statement from mathematics and negating it then saying something else follows from it (in this case nonsense). We can ask why it seems like our first set (1, 2, 5 and 6) seems reasonable, where our second set (3, 4, 7 and 8) seem not so reasonable? We may be able to say that in the first set the consequents seem to *follow* from the antecedents where as in the second set this doesn't seem to happen, it becomes hard to see what relationship there could be between the primeness of  $p$  and pigs having (or not) a pair of wings (4). This leaves us with a strange question, is a counterfactual like Edgington's

9) *If I had been 8 feet tall I would have been a freak*

Is it more like the first set (1, 2, 5 and 6) or more like the second set (3, 4, 7 and 8)? Prima facie we seem to want to say that it isn't part of the second set because it doesn't seem nonsensical like the members of the second set do. However it also doesn't seem to fit with the first set either in the sense that it is easier to deny it than any member of the first set. Even if we stipulate some sort of proviso that we could indeed go back in time and feed my past self steroids, it seems strained, we still don't seem to be able to get it to the same level of certitude as the members of the first set. Consider these other examples she provides of impossible antecedents

if I were a blackbird...

if kangaroos had no tails...

if gravity went by the inverse cube of distance...

the second is of course Lewis' own, but the third seems like a scientific statement, and what type of consequent it could have that would make it behave more like the first set of counterfactuals (1,2,5, and 6) rather than the second (3,4,7, and 8) isn't clear. Perhaps if it was something like:

*10) if gravity went by the inverse cube of distance, then objects would fall to the earth faster*

it could be part of the first set, but not if it were something like this:

*11) if gravity went by the inverse cube of distance, then I'd need to have a kidney removed by surgery*

Edgington seems to suggest that part of the problem may be that certain antecedents are now not possible to intervene on, she states “it’s not now causally possible that you become 8 feet tall. (Edgington 1995).” What this means for the study of counterfactuals is hard to say, perhaps there are different kinds of impossible antecedents, perhaps some are Necessarily impossible where as others are merely Contingently impossible. I leave this as an open question.

Many occasions exist where we do not want to assert counterfactuals with impossible antecedents but on the other hand we do not want to assert their negation either. Perhaps with regards to these counterfactuals with impossible antecedent, there is no need to give a precise account of their truth conditions as they are pointless to assert according to familiar aspects of conversational implicature that is, they are merely pointless to assert. Lewis suggests that counterfactuals with impossible antecedent should be seen as vacuously true. Given this according to Lewis we may need to reevaluate the truth conditions of our counterfactuals in such a way so that they cannot be vacuously true. This would lead to a stronger interpretation of the ‘would’ counterfactual. We could use  $[+w] \rightarrow$  to represent it and give it the following truth conditions:

$\theta [+w] \rightarrow \psi$  is true at a world  $i$  (according to a system of spheres  $\$$ ) if and only if there is some sphere  $S$  in  $\$_i$  such that  $S$  contains at least one  $\theta$ -world, and  $\theta \rightarrow \psi$  holds at every world in  $S$ .

Similarly, we also have a weaker ‘might’ counterfactual, which would be vacuously true when its antecedent is impossible. We could use  $[-m] \rightarrow$  to represent it and give it the following truth conditions:

$\theta [-m] \rightarrow \psi$  is true at a world  $i$  (according to a system of spheres  $\$$ ) if and only if every sphere  $S$  in  $\$_i$  that contains at least one  $\theta$ -world at which  $\theta \wedge \psi$  holds.

We could thus define these new strengthened and weakened counterfactual connectives in relation to each other.

- 1)  $\theta [-m] \rightarrow \psi \stackrel{\text{df}}{=} \sim (\theta [+w] \rightarrow \sim \psi)$
- 2)  $\theta [+w] \rightarrow \psi \stackrel{\text{df}}{=} \sim (\theta [-m] \rightarrow \sim \psi)$

We can get either pair from the other via the following definitions:

$$\theta [-m] \rightarrow \psi =^{\text{df}} (\theta > \diamond \theta) \rightarrow (\theta > \diamond \psi)$$

$$\theta > \square \psi =^{\text{df}} (\theta [+w] \rightarrow \theta) \rightarrow (\theta [+w] \rightarrow \psi)$$

As noted before, we considered the counterfactual with a true antecedent to be an erroneous conception, however not one that we should consider as being an automatic reason for producing a false counterfactual or a type of truth-value gap. Lewis argues that “a counterfactual with true antecedent is true if and only if the consequent is true. This is so both for ‘would’ and ‘might’ counterfactuals (and for the strong would and weak ‘might’ counterfactuals introduced in the previous section). In short counterfactuals with true antecedents reduce to material conditionals.”<sup>51</sup> Given the antecedent  $\theta$  is true at world  $i$ , it follows that there must be a  $\theta$ -permitting sphere around  $i$ , this is so because  $\{i\}$  is itself a sphere. When the consequent  $\psi$  is also true at  $i$  then it follows that there must be a  $\theta$ -permitting sphere around  $i$  in which  $\theta \rightarrow \psi$  holds that is  $\{i\}$ , therefore  $\theta > \square \psi$  is true at  $i$ . Additionally every  $\theta$ -permitting sphere contains a world where  $\theta \wedge \psi$  holds because every sphere around  $i$ , (not including the empty set which is not a  $\theta$ -permitting sphere) contains the world  $i$  itself thus it follows that  $\theta > \diamond \psi$  is true at  $i$ . Conversely, if the consequent  $\psi$  is false at  $i$ , then it follows that there can’t be a  $\theta$ -permitting sphere, thus,  $\theta > \square \psi$  is false at  $i$ . Additionally, there is a  $\theta$ -permitting sphere which contains no such world where  $\theta$  and  $\psi$  holds, that is  $\{i\}$ ; so  $\theta > \diamond \psi$  is false at  $i$ .

Counterfactuals with true antecedents have the following two valid inference patterns, the truth conditions given for the counterfactuals guarantee that if the premise is true at a world, the conclusion is true as well.

1)

$$\theta \wedge \sim \psi$$


---

$$\therefore \sim (\theta > \square \psi)$$

2)

$$\theta \wedge \psi$$


---

$$\therefore (\theta > \square \psi)$$

The first also guarantees the validity of the following two:

3)

$$\theta > \square \psi$$


---

$$\therefore \theta \rightarrow \psi$$

4)

$$\theta > \square \psi$$

$$\theta$$


---

$$\therefore \psi$$

Whether or not these consequences of the truth conditions are plausible or not is open to debate. Lewis argues that “the false information conveyed by using a counterfactual construction with a true antecedent eclipses the falsity of the truth of the conditional itself.”<sup>52</sup> If



we make a counterfactual, someone might disagree either with the antecedent or the asserted conditional itself. Lewis gives the following example.

Someone might say “If Caspar had come it would have been a good party”, another person might answer “that’s false for he did come yet it was a rotten party” or the second person might have answered “That’s true; for he did and it was a good party. You didn’t see him because you spent the whole time in the kitchen missing all the fun.” One can express agreement or disagreement with the conditional assertion and this agreement or disagreement is justified by an argument. This argument is signaled by the word ‘for’ in the replies.

These arguments have the following forms:

1)

$$\theta \wedge \sim \psi$$

---

∴ that’s false

2)

$$\theta \wedge \psi$$

---

∴ that’s true

These arguments are cogent only when the inference patterns are valid, this is seen as follows:

1)

$$\theta \wedge \sim \psi$$

---

∴  $\sim (\theta > \square \psi)$

2)

 $\theta \wedge \psi$ 


---

 $\therefore (\theta >_{\square} \psi)$ 

The truth conditions seem to be confirmed, however these may not be decisive, Lewis argues that we could counter this by saying that “it would seem very odd to pick two completely unrelated truths  $\theta$  and  $\psi$  and on the strength of their truth, to deny the counterfactual  $\theta >_{\square} \sim\psi$ ; and even odder to assert the counterfactual  $\theta >_{\square} \psi$ .”<sup>53</sup> It’s possible that the ‘that’ in the conclusions which have been judged to be true or false is not the counterfactual but rather some other belief that could be taken to be the reason for thinking the counterfactual to be true or false. Lewis asserts that counterfactuals with true antecedent and false consequent are false, this is so because no world is more similar to world  $i$ , than  $i$  itself. Additionally counterfactuals with a true antecedents and a true consequent are true, this is also the case, because no other world is even as similar to  $i$  than  $i$  itself. Lewis claims that he is more certain of the first than he is of the second. The answer Lewis thinks might be to keep the first but reject the second claim on similarity. Lewis states that “perhaps our discrimination of similarity are rather coarse, and some worlds different from  $i$  are enough like  $i$  so that such small differences as there are fail to register.”<sup>54</sup> Given this, we might try to revise the centering condition (C), which stipulated that  $\{i\}$  was to be a sphere around  $i$ . We could weaken it as follows:

Let  $\$$  be an assignment to each world  $i$  of a set  $\$i$  of sets of worlds. The  $\$$  is a weakly centered system of spheres if and only if, for each world  $i$ , the following conditions hold:

(W)  $S_i$  is weakly centered on  $i$ ; that is  $i$  belongs to every non empty sphere around  $i$ , and there is at least one non empty sphere around  $i$ .

(1)-(3)  $S_i$  is nested, closed under unions and closed under (non empty) intersections; these conditions are unchanged.

### Inner and Outer Modalities

In this weakly centered system of spheres, the inner most non empty sphere around  $i$  is the intersection of all non empty spheres around  $i$ , that is,  $\bigcap(S_i - \{\Lambda\})$ , this yields worlds differing only negligibly from  $i$ , so that these worlds will then come out just as close to  $i$  as  $i$  itself. Weakening the centering condition leaves the truth conditions of counterfactuals unchanged and the result is the desired one. Counterfactuals with true antecedents and false consequents must be false. Counterfactuals with true antecedents and true consequents can come out to be either true or false. A false result for a counterfactual with a true antecedent and true consequent can be achieved in the following way. If we allow for a sentence  $\theta$  to be true at a world  $i$ , then it follows that the smallest  $\theta$ -permitting sphere around  $i$  will be the innermost non empty sphere around  $i$ . This sphere will contain  $i$  itself. This sphere may also contain (or it may not) other worlds. When it does there may or may not be  $\theta$ -worlds other than  $i$  among them. If there are, then  $\theta >_{\square} \psi$  holds at  $i$  if and only if the consequent  $\psi$  holds not just at  $i$  itself but also at

the other  $\theta$ -worlds in the innermost sphere around  $i$ . Given this it follows that in such a way “it may happen that a counterfactual with true antecedent and consequent is false if the consequent is false at a sufficiently close antecedent world.”<sup>55</sup> By weakening the centering condition a distinction arises then between “truth at  $i$  itself and truth at all or some of the worlds in the innermost non-empty sphere around  $i$ .”<sup>56</sup> This will then give rise to the notion of inner modality, that is inner necessity and inner possibility. We can define these in the following ways.

$$*\square\theta =^{\text{df}} \top [+w] \rightarrow \theta \text{ or } (\diamond\top \wedge \top > \square\theta)$$

and

$$*\diamond\theta =^{\text{df}} \top [-m] \rightarrow \theta \text{ or } (\square\top \rightarrow \top > \diamond\theta)$$

We can give these new modal operators the following truth conditions:

$*\square\theta$  is true at  $i$  if and only if  $\theta$  is true at every world in some non-empty sphere around  $i$ .

$*\diamond\theta$  is true at  $i$  if and only if  $\theta$  is true at some world in every non empty sphere around  $i$ .

The inner modalities are defined in terms of accessibility, which is the assignment to each world  $i$  of the innermost non-empty sphere around  $i$ , that is  $\cap(\mathcal{S}_i - \{\Lambda\})$ , as its single sphere of accessibility. We can say for  $*\square\theta$  to be true,  $\theta$  must hold at every maximally close world and likewise for  $*\diamond\theta$  to be true  $\theta$  must hold at some maximally close world. Lewis asserts that “the outermost sphere includes the innermost nonempty sphere; therefore outer necessity is stricter than inner necessity. Therefore  $\square\theta$  implies  $*\square\theta$  and  $*\diamond\theta$  implies  $\square\theta$ .”<sup>57</sup> In our weakly centered

system of spheres the inner modalities can be defined more simply as  $\top >\Box \theta$  and  $\top >\Box \theta$  respectively. Lewis holds that despite these considerations the centered system of spheres is likely more appropriate for the interpretation of counterfactuals, consequently in a centered system of spheres it becomes rather pointless to consider the inner modalities, in such a case  $*\Box\theta$  and  $*\Diamond\theta$  can be said to be equivalent to  $\theta$  itself.

### Counterfactual Fallacies

We will now move to examine some counterfactual fallacies. Lewis claims that “certain inferences are correct for the material conditional, and indeed any constantly strict conditional, but not for variably strict conditionals. The inference fails because the strictness varies between conditionals in the premises and conclusion.”<sup>58</sup> There are three fallacies of interest, these are the fallacy of strengthening the antecedent, the fallacy of transitivity and the fallacy of contraposition.

*The fallacy of strengthening the antecedent.*

This fallacy is the following invalid inference pattern:

$$\theta >\Box\psi$$


---


$$\therefore \theta \wedge \chi >\Box\psi$$

Lewis generalizes this inference pattern in the following way:

$$\Box (\chi \rightarrow \theta)$$

$$\theta > \Box \psi$$


---


$$\therefore \chi > \Box \psi$$

Lewis instantiates this argument in the following way:

$$\Box (\text{I started at 5 this morning} \rightarrow \text{I started before 6})$$

If I had started before 6, I would have arrived before noon.

---

$\therefore$  If I had started at 5, I would have arrived before noon.

Obviously the first premise is true. The second premise can be true yet the conclusion might still turn out to be false. If we suppose that he started just before 6 and tried a new shortcut that cut 2 hours off the trip and thus arrived at noon exactly. But if we suppose that he started at 5, he would have been too sleepy to remember to try the shortcut. Lewis is supposing that the later the person starts in the range of times permitted by the antecedent, the closer antecedent-world will be to our world. This might be the case but equally it might not. He may have planned have planned on starting at 5 but failed because his alarm did not wake him. The inference is fallacious even when outer necessity is logical necessity and it follows that this is also the case when less strict necessity is applied. Jeffrey Dunn has argued that “inferring  $((p \wedge q) > r)$  from  $(p > r)$  is invalid. But when we evaluate counterfactuals in everyday life, we often must do something very much like strengthening the antecedent. If I hadn’t set my alarm last night, then

what would have happened? Without some way to strengthen the antecedent—to indicate that we don't stray too far from actuality—it's far too indeterminate to say”<sup>59</sup>

Another perhaps slightly easier to follow example of strengthening the antecedent for counterfactuals might look something like the following:

If the Conservative party had not won the last election, then the Liberal Party won it.

---

Therefore, if the Conservative party had not won the last election and the New Democrat Party had received 90% of the popular vote, then the Liberal Party would have won the last election.

Obviously this is a poor argument. If the New Democrat Party had received 90% of the popular vote they would have won the election. The account of counterfactuals which Lewis gives explains this quite well. The truth of  $\theta >_{\square} \psi$  requires that the nearest  $\theta$ -world be a  $\psi$ -world but the nearest  $\theta$ -world might not be an  $\psi$ -world. To find the nearest world that is  $\theta$  and  $\psi$  we might have to move to a still more distant world and that world clearly might not be a  $\psi$ -world.

### *The fallacy of transitivity*

This fallacy is the following invalid inference pattern:

$$\chi >_{\square} \theta$$

$$\theta >_{\square} \psi$$


---


$$\therefore \chi >_{\square} \psi$$

Lewis argues that the fallacy of transitivity is a further generalization of the fallacy of strengthening.

We can instantiate this argument in the following way:

If John had attended the concert then Allan would have attended the concert

If Allan had attended the concert then Tim would have attended the concert

---

Therefore if John had attended the concert, then Tim would have attended the concert

Another, more famous example of the fallacy of transitivity for counterfactuals comes from Robert Stalnaker, this is of course the J. Edgar Hoover example and it is as follows:

If J. Edgar Hoover had been born a Russian, then he would have been Communist

If he had been a Communist, he would have been a traitor.

---

Therefore if he had been born a Russian, he would have been a traitor

This fallacy is more obviously faulty than that of strengthening the antecedent. As Lewis argues in this case “the antecedent of the first premise must be more far fetched than the antecedent of the second, which is the consequent of the first. Then the closest worlds where the first antecedent holds are different from and may differ in character from the closest worlds where the



second antecedent holds.”<sup>60</sup> Interestingly the view that this argument form of transitivity for counterfactual is invalid has not been fully accepted by everyone. Indeed, most notably E.J. Lowe has called into question whether in fact the J. Edgar Hoover and other such like examples can be taken seriously. According to his analysis “all the worlds in which Hoover is born in the Soviet Union, including the closest such world – which is a world in which he is a communist are less close than the closest world in which Hoover is a communist; so that even though he is a traitor in the latter world for it is a world in which he is a United States citizen holding a high public office that is incompatible with allegiance to communism – it need not follow that he is a traitor in the closest world in which he is born in the Soviet Union.”<sup>61</sup> The problem according to E.J. Lowe with the J. Edgar Hoover example and others like it, resides in the seeming ambiguity, it represents according to Lowe the fallacy of *equivocation*, the consequent in the first premise does not truly express the same proposition as the antecedent in the second premise they represent different abbreviations or ellipses of different sentences. In this way Lowe seems to be in line with Crispin Wright’s thinking with regards to the same possible world being used in a single piece of reasoning (Wright 1983 and Lowe 1990). By the first sentence we mean something like ‘if J. Edgar Hoover had been born in the Soviet Union then he would have been a Soviet citizen and a communist’. By the second we mean something like this ‘if J. Edgar Hoover had been a communist and had still been a citizen of the United States then he would have been a traitor’. If these accounts of the premises are correct then it shows that transitivity is not really invalid as an inference form for counterfactuals because the J. Edgar Hoover example does not really represent the form specified by transitivity. In some senses this interpretation may be somewhat ad hoc as other examples may not be explained away so easily. Lowe has stressed that when we are presented with counter examples of transitivity we are only able to judge the premises as

being true and the conclusion as false if we adopt a shifting standard of similarity between possible worlds for the purposes of evaluating the counterfactuals, a shifting standard is tantamount to a kind of equivocation.

According to Lowe if we examine the J. Edgar Hoover example carefully and if “we are to judge the premises of this argument to be true and its conclusion to be false, must we not vary our assumptions as to which features of a possible world are more important for the purpose of assessing its degree of overall similarity to the actual world.”<sup>62</sup> It’s clear that for almost any counterfactual conditional it is clearly possible to imagine some conversational context in which it would be natural to judge that particular conditional as being true. Even if this is possible for a single statement it becomes questionable whether this is possible to do when we take two or more related counterfactuals. As Lowe puts it concerning the J. Edgar Hoover example, we “cannot – I suggest envisage a single conversational context in which it would be natural to judge its premises to be true and its conclusion to be false. If that is so, then the Hoover arguments, construed as having true premise and a false conclusion does indeed involve a kind of equivocation. And that means that it cannot fairly be represented as constituting a genuine counter example to (Trans).”<sup>63</sup> Lowe suggests his own truth condition for the counterfactual statement which will in his view preserve the inference of transitivity.

(L) ‘If it had been the case that  $P$  then it would have been the case that  $Q$ ’ is true if and only if (a) in every sufficiently close possible world  $P \rightarrow Q$  is true, (b) either there is some sufficiently close possible world in which  $P$  is true or else in every sufficiently close possible world  $Q$  is true.

Clause (a) of (L) requires that the counterfactual imply the corresponding material conditional and clause (b) requires that the antecedent be a possible truth or else that its consequence be a necessary truth. Analysis (L) implies that transitivity is valid and that neither contraposition nor strengthening is valid. Lowe contends that his analysis of counterfactuals supports the distinction between ‘would’ and ‘might’ counterfactuals however it does not support the principle of Conditional Exclude Middle (the principle can be stated as  $(P >_{\square} Q) \wedge (P >_{\square} \sim Q)$  and is meant as way of making any system which rejects one of the disjuncts accept the other) Furthermore Lowe argues that any account of counterfactuals which employs a relationship of ‘closeness’ of possible worlds must in some sense accommodate the context dependency of the propositional content of counterfactuals.

In all these examples the ‘might’ counterfactual  $\theta >_{\diamond} \sim \psi$  is true. In some but not all cases the ‘would’ counterfactual  $\theta >_{\square} \sim \psi$  is non vacuously true. These inference patterns turn out to be true when a third premise is added.

$$\sim (\theta >_{\square} \sim \chi)$$

$$\chi >_{\square} \theta$$

$$\theta >_{\square} \psi$$

---

$$\therefore \chi >_{\square} \psi$$

or, more simply:

$$\theta >_{\square} \chi$$

$$\chi > \Box \theta$$

$$\theta > \Box \psi$$

---

$$\therefore \chi > \Box \psi$$

another special case of transitivity is also valid

$$\chi > \Box \chi \wedge \theta$$

$$\chi \wedge \theta > \Box \psi$$

---

$$\therefore \chi > \Box \psi$$

or, more simply:

$$\chi > \Box \theta$$

$$\chi \wedge \theta > \Box \psi$$

---

$$\therefore \chi > \Box \psi$$

Another valid inference pattern, which might be called ‘weakening the consequent’ is as follows:

$$\chi > \Box \theta$$

$$\Box(\theta \rightarrow \psi)$$

---

$$\therefore \chi > \Box \psi$$

*The Fallacy of contraposition.*

This fallacy is the following invalid inference patterns:

$$\theta > \Box \psi$$

---

$$\therefore \sim \theta > \Box \sim \psi$$

and

$$\sim \theta > \Box \sim \psi$$

---

$$\therefore \theta > \Box \psi$$

Clearly, either both are valid or neither are valid. An instantiation of the first would look something like this:

If Boris had gone to the party,  
Olga would had still have gone.

---

Therefore if Boris had not gone,  
Olga would still not have gone.

If we suppose that Boris wanted to go but had stayed away in order to avoid Olga, thus the conclusion is false. However Olga would have gone even more willingly if Boris had been there so the premise is true. However  $\theta >_{\square} \psi$  can be true even if its contrapositive isn't.

Another example of this argument form could be seen in the following example:

If John had moved into the house, then Jane would not have moved out.

---

Therefore if Jane had moved out of the house, then John would not have moved in.

It's easy again to see how this argument form is invalid. It is reasonable to imagine a state of affairs that makes the premise true and the conclusion false. If we suppose that Jane wanted to live in the same house as John, however the feeling is not reciprocated by John. Had John moved into the house which Jane lived in, Jane would have been happy with the result and would have wished to remain there. John had wanted to move into the house, however he was dissuaded from doing so based on Jane's being present. In this way we see that the conclusion is false while the premise is true. We can explain this situation on Lewis' account when we see that  $\theta >_{\square} \psi$  requires that the nearest  $\theta$ -world be a  $\psi$ -world. If the nearest  $\psi$ -world were the nearest  $\theta$ -world then it obviously follows that  $\sim \psi >_{\square} \sim \theta$ . However it need not be the case that it is, the nearest  $\sim \psi$ -world could be further away, in such a case further away than the nearest  $\theta$ -world. Then it would not follow that such a world must be a  $\sim \theta$ -world.

Can use contraposition to justify the following inference pattern involving ‘might’ counterfactuals?

$$\theta > \diamond \psi$$

---

$$\therefore \psi > \diamond \theta$$

Unfortunately not, his inference pattern for might counterfactuals has no plausibility whatsoever. Contraposition of counterfactuals may be invalid, however inference by modus tollens on a counterfactual is valid.

$$\theta > \square \psi$$

$$\sim \psi$$

---

$$\therefore \sim \theta$$

It should be stressed however that this is not modus tollens as proceeding from contraposition followed by modus ponens as in the case for material conditionals. Rather for the counterfactual it is an inference from a counterfactual to a material conditional followed by contraposition on the material conditional followed by modus ponens on the contraposed material conditional.

We have now completed our evaluation of the technical aspects concerning Lewis's system of counterfactuals. We examined the basic elements of Lewis's system and found that the strict conditional was not appropriate for a theory of counterfactuals rather a variably strict conditional was better suited. We took a look at the centered system of spheres elucidated by Lewis in detail paying special attention to the four conditions and their inter relationships. We studied the four truth conditions for counterfactuals and found that we can identify two cases where the 'would' counterfactual is true and two where it is false. We evaluated the Limit assumption and saw that despite being problematic in some senses it was necessary for the system to function. The modal interpretations of the system were considered as well, along with the issues of impossible antecedents. Finally we analyzed the notion of inner and outer modalities in the centered system of spheres and the counterfactual fallacies. Up until this point we have made mention of the possible worlds without characterizing them too deeply, in the next section we will attempt to give some color to this underlying concept.



### **Chapter 3: The Nature of Other Worlds, Concrete or Abstract?**

The other day I went to the store to purchase a shirt. I was faced with the choice between two shirts of the same style, one red shirt and one blue shirt. I tried both on and although I liked the way they both looked in the end I decided to buy the blue one. That evening I went to my sister's for dinner and I wore the new shirt. My sister commented that she liked the shirt I was wearing. I mentioned to her that I had decided to take this one instead of the red one. She paused for second, and then said that I should have got the red one instead because it would have looked better with the pants I was wearing. I went back to the store the next day to see about the red shirt, however it was gone and when I asked the store clerk about it he said someone else had come in just after me and purchased it, I asked if he would be getting anymore of those shirts, he said that they would not be restocking those particular shirts, as it was the end of the season. I cursed my luck and pondered my fate of being stuck with only the blue shirt. This story represents a classic divergence between two possible states of affairs, in this case my having the blue shirt or my having the red shirt. Although it's clear that I exist in a world where I own the blue shirt, I could just have easily chosen the red one, in fact it was quite a haphazard choice between the two.

The question arises then if I could just as easily have picked the red one is there then a world in which I did? We can answer this question in several different ways. One way is to say that there is a theoretical state of affairs that we can conceptualize in abstract terms, in which an agent who takes my name and stands for me, purchases a red shirt. Another way we could

approach this is to say that there is parallel world as concrete and 'real' (as unhelpful as the term might be in this case) as the world we live in, where a person like me does indeed own the red shirt (perhaps he is now cursing his luck and wishing that he had purchased the blue one). A third answer, is to say that there is no such other world, either abstractly or in concrete terms which exists, I'm stuck with the blue shirt and I have to live with it. The first type of answer seems to cost us little or nothing in the way of ontology, we need only create an abstract construct in a system to account for the other possibility. However the second answer is much more expensive in terms of ontology, the reason for this is that it forces us to give credence to a plethora of entities for every possible divergence that presents itself in our world. The third answer is taken to be the most radically austere, as it gives ascent to neither an abstract nor a concrete rendering of the possibility of my owning the other shirt and is felt by most philosophers to be an unacceptable and defeatist way of dealing with possibility.

The second answer is of course the answer that David Lewis puts forward in his doctrine of Modal Realism as argued for in both *Counterfactuals* and *On the Plurality of Worlds*. Lewis holds that Possible Worlds (including our own) are real concrete things which are isolated from one another both spatially and temporally, but also causally. The other worlds are not different from our world in kind, they have the same type of existence as our own world. According to Lewis' nothing special differentiates our world from the other worlds except that we happen to live in it. In this way, our world is 'actual' but only in terms of us, from our point of view, it of course stands to reason that the inhabitants of the other worlds are similarly inclined to view their own world as the actual world. Lewis argues that his plurality of worlds aids us in our accounts concerning modality. Modal Realism according to him can be used to found our analysis of modal logic and consequently, with that, our account of counterfactuals. For him modality is

then quantification over possible worlds and counterfactuals are statements of fact about the various possible worlds. Lewis argues that the appeal of Modal Realism is its utility and it is analogous to set theory in mathematics, in this way he argues that the cost of his Modal Realism in terms of ontology is well worth the expense as it gives us the means to approach many problems in Philosophy and Science and makes them much more tractable. He argues that the price is right as in mathematics, even if the outcomes aren't quite as spectacular.

The following question arises, 'is this really tenable?' Can we really give ascent to these Possible Worlds as full blown concrete entities on par with the actual world? Two issues are at play here, the first is the utility and fruitfulness of the thesis of Modal Realism that is the concrete existence of these alternate possible worlds, the second is the evidence for their existence, in other words what proof can we have for their existence. It's important to keep in mind that these two issues are somewhat separate; we might (as we shall see) find great use for the possible worlds in terms of the avenues they open in our intellectual pursuits, while not being able to conclusively establish their existence as genuine entities with full fledged concrete existence. It should be stressed that many theoretical entities can be posited and indeed bring fruitful results however we might have no good reason to believe that they actually exist, they might just be empty theoretical entities in our theories. In the language of Ian Hacking we should ask 'can we spray them?' that is to say can we have any manipulative or experimental scope over these alleged concrete other worlds? The answer to this may turn out to be 'no'.

In view of this, we will attempt to evaluate Lewis's doctrine of Modal Realism. To do this we will sketch the position as presented in both *Counterfactuals* and *On The Plurality of Worlds*. Our focus will be on characterizing the position in detail and then we will briefly

evaluate the rival account of possible worlds as abstract constructs, known as Ersatz Modal Realism. Finally we will try to take stock of where all of this may leave us.

Lewis advocates the thesis of Modal Realism, that is the contention that the world we inhabit is but one among many worlds. The other worlds are not in any way spatially or temporally related to our world. No distance or amount of time separates our world, the actual world, from any of the other possible worlds. Additionally each world, including our own, is isolated from each other world. There are no spatiotemporal relations of any kind between things which exist in different worlds. Each world is causally closed unto itself and nothing which occurs at one world causes anything to occur at another. No world has a part or parts in common with any other world, that is to say there is no overlap between worlds.

Lewis argues for what could be called a robust and rich ‘pluriverse’ which encompasses the full scope and domain of that which is possible. In terms of the number and amount of other worlds he argues that “there are so many other worlds, in fact absolutely every way that a world could possibly be is a way that some world is.”<sup>64</sup> So from the most pertinent of details like the outcomes of elections and great battles to the most trivial accidents like the choice between combinations at Chinese restaurants made by diners, a separate and hermetically closed world exists. Of course, this only commits Lewis to accept a plurality of things which most of us already accept. Most of us would be hard pressed to deny that there does not in fact exist an actual world (skeptics notwithstanding), Lewis of course accepts this as well, however he commits himself to an ontology of infinite entities exactly like the actual world. So as Lewis argues, when he posits the existence of the other worlds he is by no means asking us to believe in any new objects merely asking us to believe in a collection of things of which we already accept the existence of at least one of them. This, Lewis argues may be unparsimonious, but not so in

an unforgivable way. According to him there are two distinct types of parsimony, one qualitative, one quantitative. Qualitative parsimony concerns the number of different kinds of objects that a theory posits. Thus, any theory that posits both bodies and spirits, is less qualitatively parsimonious than a theory which posits only bodies or spirits, but not both. Quantitative parsimony on the other hand concerns the number of instance of a particular kind of object that has been posited. Thus if we believe in something like Universal Laws of Nature, then an account of these laws that has 100 rather than 1000 laws is more quantitatively parsimonious. Lewis argues that his “realism about possible worlds is merely quantitatively not qualitatively unparsimonious. You believe in our actual world already I ask you to believe in more things of that kind, not in things of some new kind.”<sup>65</sup> Further the “other worlds are of a kind with this world of ours. To be sure, there are difference of kind between things that are parts of different worlds – one world has electrons and another has none, one has spirits and another has none – but these differences of kind are no more than sometimes arise between things that are parts of one single world, for instance in a world where electrons coexist the difference between this and the other worlds is not a categorical difference.”<sup>66</sup> Whether or not this type of parsimony can be accepted by the ontologically skittish amongst us is open to debate. Many have argued, most notably Quine; that this expansive account of possible worlds is metaphysically disorderly and thus doesn’t warrant our allegiance.

Our world is no different from the others in its manner of existing, that is to say that the other worlds have the same existential status as our own world. They are no more or less real than our own world. What is more, it follows then that we should not consider the other worlds, according to Lewis, as our creation, they have an independent existence apart from our theorizing and would have existed even if they had never entered into our considerations. If this

is the case, then we might ask what on earth are these possible worlds for? Lewis argues that the case for possible worlds is buttressed by the fact that they are very useful to posit with regards to various fields of intellectual endeavor. It's hard to disagree with this given that possible worlds have been unquestionably helpful with regards to logic, philosophy of mind and language, science and metaphysics. Lewis argues that "because the hypothesis is serviceable" it becomes clear that it "is a reason to think that it is true."<sup>67</sup> This is, at least in some senses, a type of abduction, or inference to the best explanation. We can believe in the theory of an infinite pluriverse of alternate possible worlds precisely because it has been so very useful up until now where we have applied it. How could it be so useful yet false? At least this is the contention of the Modal Realist. It should be stressed however that many things have been and indeed can be very useful yet false, as noted earlier the existence of something and its usefulness in our intellectual pursuits are separate issues requiring separate treatment.

Lewis draws the parallel between his doctrine of possible worlds and that of set theory in mathematics. Indeed the title of his the first chapter in *On the Plurality of Worlds* is 'The Philosophers Paradise' which is clearly a reference to the so called 'Cantorian Paradise' of set theory. The benefits of the expanded set theory purposed by Cantor are indeed well worth the expense in ontology, but this just might be due to the fact that the expense is very low. Lewis draws the parallel in the following way "as the realm of sets is for mathematicians so logical space is for philosophers. We have only to believe in the vast realm of possibilia and there we find what we need to advance our endeavors."<sup>68</sup> It costs very little to expand the universe of discourse in set theory to the dimensions required by Cantor. Not so for Lewis's Modal Realism, contemporaries of Lewis have noted this. Allen Stairs, has argued that the "extra entities we accept when we embrace full set theory are all constructions out of entities we already took

ourselves to have on hand. They are parasitic upon those entities in the clear sense that ultimately, they decompose into them. Nothing of this sort is true of other worlds. They are emphatically not construction out of entities that we already took to be unproblematic. That is why the principle of recombination is not really like the principles for constructing sets.”<sup>69</sup> Notwithstanding this, Lewis holds that the price is indeed right, because the “benefits of theoretical unity and economy are well worth the entities” and further it is “worth believing in vast realms of controversial entities for the sake of enough benefit in unity and economy of theory.”<sup>70</sup> The hope then is that if the theoretical benefits are to be had by an expanded ontology of possible worlds then their use will be justified in the various fields they have been employed in.

We will now proceed to evaluate the doctrine of Modal Realism in terms of the aspects that Lewis raises, that is, isolation, concreteness, plenitude and actuality. We will evaluate each one in turn.

### **Isolation**

With regards to isolation as previously noted, each world is both spatiotemporally and causally separated from every other world that exists. We can say that each world has its proper parts and that it shares none of these with any other world. The parts of these worlds, or more properly the individuals, are called by Lewis ‘worldmates’. The world is then the combination of all the individuals that make it up; that is to say that any given world is the logical sum of all the possible individuals in it. It follows from this that a world is a maximal sum, that is to say

that anything which is a worldmate of anything else is itself a part of the world in question. Lewis argues that “whenever two possible individuals are spatiotemporally related they are worldmates. If there is any distance between them – be it great or small, spatial or temporal – they are parts of one single world.”<sup>71</sup> In other words given any two possible individuals, if every part of one is spatiotemporally related to every part of the other individual and they are distinct objects then it follows that they must stand in the relation of worldmates to each other. If two objects in two separate worlds bear the same relations to similar objects in their respective worlds then these two objects cannot be worldmates, however, in Lewis’s language they will be *counterparts* to each other across possible worlds. Lewis puts it as follows “there are counterpart centuries, or weeks or seconds. Like wise there are counterpart places: galaxies, planets, towns. So things that are parts of the two worlds may be near or far from one another in very natural counterpart-theoretic senses”<sup>72</sup> The worlds are unified wholes constituted and defined by the inter relations between their parts in spatiotemporal terms. Clearly there can be no such relations between objects across worlds and this then provides us with a clear way of drawing the boundary between worlds. Lewis describes this as follows, he argues that “each world is interrelated (and is maximal with respect to such interrelations) by a system of relations which if they are not spatiotemporal relations rightly so called are at any rate analogous to them.” It of course follow that “any natural *external* relation will do to unify a world. Every part of a world bears some such relation to every other part but no part of one world ever bears any such relation to any part of another.”<sup>73</sup> Another way of putting this is to say that causal closure in a world must hold.

With regards to causation clearly only individuals within the same world can stand in the relations cause and effect to one another. The result of this is that all worlds lie in total causal



isolation from one another and this is therefore a principal means of demarcation between worlds. Lewis can account for cause in counterfactual terms by the following means. Given any world  $W$  where an event  $c$  is the cause of an event  $e$ . Both the cause and the effect are distinct and unique events. If it is the case at  $W$  that  $c$  does not occur then  $e$  will not occur. We can interpret this in terms of possible worlds by saying that at the closest possible worlds to  $W$  where the cause does not occur the effect does not occur either. Lewis goes on to explain precisely why any type of trans-world causation is strictly speaking impossible. If we take a putative example of trans-world causation we can see that “event  $c$  occurs at world  $W_c$ , event  $e$  occurs at world  $E_e$ , they are distinct events if  $c$  had not occurred,  $e$  would not have occurred either. This counterfactual is supposed to hold – where? It means that at the closest to – where? At which  $c$  does not occur  $e$  does not occur – where? Either.”<sup>74</sup> If the counterfactual is meant to hold at the world where  $c$  causes  $e$  then perhaps it should hold at both these two separate worlds. However this leads to contradictory results that can’t be accepted despite Lewis’s best efforts to find a way. No matter which world we remove the cause from, we don’t seem to get the result we want. If  $c$  and  $e$  truly are independent events in separate worlds it’s hard to see how eliminating  $c$  from  $W_c$  will have any effect on  $e$  on at  $E_e$ . As Lewis puts it “there isn’t any trans-world causation. And not because I so stipulated as a principle of demarcation, but as a consequence of my analysis of causation and counterfactuals,”<sup>75</sup> interestingly this points the way to an answer to one of the most interesting questions we can ask about the other worlds, why is it that we can’t travel between the worlds? Lewis answers this by stating that “if there is no trans-world causation, there is no trans-world travel. You can’t get into a ‘logical-space ship’ and visit another possible world” it follows from this that if there is no trans-world causation then there is

no trans-world causal continuity, if there is no causal continuity then there is no survival and if there is no survival there can be no travel.<sup>76</sup>

### **Concreteness**

Lewis' account of possible worlds as full blown existents separate from this world encounters a deep and troubling problem, as Allen Stairs has noted "if other worlds are like many worlds of the eponymous interpretation of quantum mechanics, or for that matter, like Afghanistan or the center of the sun – if they are concrete, physical things – then they require the same sort of evidence that these other things require. This evidence is ultimately based on the existence of causal chains that connect us to the things in question. Since other worlds are causally isolated from us, we can't have this kind of evidence, and so we can have no reason for believing in them."<sup>77</sup> This is the same claim that Brian Skyrms makes, he contends that if we are to take the other worlds as concrete in the same way as other familiar concrete objects, "they require the same sort of evidence for their existence as other constituents of physical reality" and clearly what is "required to show that the sort of possible worlds Lewis wants exist, is their presumption in the best physical theory."<sup>78</sup> This objection seems deeply problematic for Lewis, in fact as seen, Lewis doesn't deny it, instead he supports his position by arguing that the utility of the assumption of concrete alternative possible worlds more than compensates for our inability to establish their existence empirically. We will now move to evaluate some of the issues with regards to Lewis's treatment of concreteness as they apply to the possible worlds of his Model Realism.

Lewis seems at some times to equivocate on the precise nature of the other worlds. Despite arguing that the other worlds are exactly like our own, which seems to suggest that they are concrete, he however appears to suggest otherwise, he argues that “the other worlds are of a kind with this world of ours, doubtless you will expect me to say that possible worlds and individuals are concrete not abstract. But I am reluctant to say that outright. Not because I hold the opposite view; but because it is not clear to me what philosophers mean when they speak of ‘concrete’ and ‘abstract’ in this connection.”<sup>79</sup> This seems strange because under no circumstances would Lewis disagree with the assertion that this world of ours is indeed concrete, so for the other worlds to be like this world they indeed must be concrete. So it seems despite his reservation Lewis is forced to admit that the other worlds must be concrete entities. He seems to do just this arguing “even without knowing what concrete is supposed to mean. I take it at least, that donkeys and protons and puddles are supposed to be paradigmatically concrete.”<sup>80</sup> According to his Modal Realism then, each one of these objects which are parts of our world, which are certainly concrete entities, has a duplicate in the other worlds, which are attributed as parts of these other worlds. As we have previously seen other worlds and their parts can in no way be related in spatiotemporal or causal terms to this world and its parts or any other world and its parts. This isolation from one another makes it obvious that if this wasn’t so, that is, if there were any type of spatiotemporal or causal relation between two parts, then these two parts would necessarily be parts of a larger world.

Lewis argues we should resist the temptation to see the other worlds as less concrete and more abstract because this becomes unpalatable for obvious reasons. We simply can’t make the abstract/concrete distinction a relative matter and remain consistent. It just will not do to say that

for us the other worldly individuals are abstract and for them we are abstract. As Lewis argues “one thing is certain: whatever the abstract-concrete distinction is, at least it’s supposed to be very fundamental difference between two kinds entities. It has no business being a symmetrical and relative affair”<sup>81</sup> So it won’t help on Lewis’ account to either consider our own world or the other worlds as abstract entities. Whether or not certain parts of a given world can be abstract entities is an open question that Lewis is not prepared to answer either in the affirmative or the negative. Lewis concludes by finding that “the worlds are concrete. They lack no specificity, and there is nothing for them to be abstract from. As for the parts of worlds, certainly some of them are concrete, such as the other-worldly donkeys and protons and puddles and stars.”<sup>82</sup> This of course leaves open the door to the possibility that there might be some worlds which have universals and abstract entities in them along with their concrete ones.

### **Plenitude**

We will now move to evaluate the issues Lewis raises regarding the plenitude of worlds. At its core the issue of plenitude revolves around the abundance of worlds, that is to say, there must be a sufficient worlds to instantiate all the possibilities that there are. There must be no unfulfilled possibilities; that is, gaps or omissions in the logical space which the possible worlds doctrine provides. Lewis tentatively advances the two following principles to express the notion of plenitude:

(1) Absolutely every way that world could possibly be is a way that some world is.

(2) Absolutely every way that a part of a world could possibly be is a way that some part of some world is.

Lewis argues that it makes sense to equate the ways that the world can be with the worlds themselves if only in terms of economy of the theoretical apparatus. However the result of this as Lewis indicates is that (1) will be without content. It turns out to be true for any number of worlds and gives us nothing in the way of the plenitude of worlds that we seek. The same can be said for the second principle. The possibility remains open to equate the ways a world could be with a set, either the unit set or an equivalence set, but this seems unsatisfactory and does nothing to give the principles content. Perhaps there is no hope then for (1) and (2), Lewis concludes just that. Instead of trying to amend or salvage the two principles Lewis suggests “that we look to the Humean denial of necessary connections between distinct existences to express the plenitude of possible worlds. I require a *principle of recombination* according to which patching together parts of different possible worlds yields another possible world. Roughly speaking, the principal is that anything can coexist with anything else at least provided they occupy distinct spatiotemporal positions. Likewise anything can fail to coexist with anything else.”<sup>83</sup> It might be advisable to resist such a formulation as it stands, particularly the notion that anything should be permitted to coexist with anything else. The reason for this according to Lewis is that individuals and objects really do only exist at one world uniquely (though suitable copies may exist in other worlds) there is no overlap between worlds thus an object can only reside in one world at a time. Lewis argues that normally we could invoke counterpart relations as a substitute for trans-world identity but in this case it won’t do. It seems implausible to suggest that any counterpart of anything can coexist with the counterpart of

anything else. The relation of counterpart is above all underlined by the notion of similarity, most importantly similarity of origin counts as the decisive factor. Given this we don't speak of 'counterparts' rather we can speak of 'duplicates' of objects and individuals whose origin is similar enough to fall into this relation. It should be clear that the notions of counterpart and duplicate are vital to Lewis' argument for the existence of other worlds as concrete entities.

Another issue raised by Lewis concerns the number of individuals that can coexist with one another in the same world. If there can be a possible individual then it stands to reason that there can be two of the same individual coexisting in the same world, thus it follows that there could be any number of coexisting duplicates even an infinite amount. However there is an upper limit on the number of duplicates even if we allow for it to be infinite. As Lewis puts it only "a limited number of distinct things can coexist in the space time continuum. I cannot exceed the infinite cardinal number of the points in a continuum. So if we have more than a continuum many possible individuals to be copied, or if we want more than continuum many copies of any single individual, then a continuum will be too small to hold all coexisting things that our principle seems to require."<sup>84</sup> This principle of recombination might lead to the conclusion that space time is larger than initially expected but Lewis cautions against this as it would be a consequence that seems implausible given that the goal was to map out the dimensions of possible space-time. Of course there will be ultimately the upper limit of what can exist in a space-time given that it must have a size and shape, beyond this, Lewis can have any two objects exist together and any number of duplicates.

The suggestion was raised by Quine that possible worlds might be treated as Mathematical representations. The representations would be formulated in terms of a set of coordinates that appear as an ordered quadruple which would correspond to the space-time

points which can be occupied by matter. We assume for the sake of simplicity that the worlds in question are confined to Euclidian four dimensional space. The worlds then can be represented as a set of four real numbers 0 and 1, of the form  $\langle x, y, z, t \rangle$ . The first three of the numbers correspond to spatial arrangements and fourth to a temporal one. From this, we create equivalence classes of these sets which represents the possible distribution of matter. This of course is a simplification, but it serves the intended purpose of rendering clear, what it would mean to give a more precise representation to the possible worlds. Of course it stands to reason that for Lewis this mathematical representation will be just that, a representation and cannot be a replacement for the worlds themselves. The reason for this is that this would tantamount to endorsing an ersatz conception of possible worlds; that is, that the worlds are purely abstract constructions or representations contrary to Lewis's position that the worlds be taken as concrete entities. Lewis argues that "we should not identify the worlds with any such mathematical representations,"<sup>85</sup> but that doesn't mean that the representation is without any use whatsoever, on the contrary, this mathematical representation provides for an excellent way of keeping track of the worlds. We will have more to say about this shortly.

According to Lewis, there are compelling reasons however to resist the temptation to equate the worlds with mathematical representations the reason for this is that if we follow this line of reasoning we might be forced to accept that the actual world is a mathematical representation, this is a conclusion few would be willing to accept. Lewis states "I cannot believe (though I don't know why not) that our world is a purely mathematical entity. Since I do not believe that other worlds are different in kind from ours, I do not believe that they are either. What is interesting is not the reduction of worlds to mathematical entities, but rather the claim that the possible worlds stand in a certain one-to-one correspondence with certain mathematical

entities.”<sup>86</sup> The situation is analogous to the one we treated with regards to considering the worlds as abstract, as we saw before if we hold that the other worlds are abstract entities and following Lewis, that this world isn’t really different in kind from the other worlds. We seem forced, under these circumstances, to treat the actual world as an abstract entity, which again few people, (barring perhaps radical Platonists) would be willing to accept. Lewis claims that “we should accept a *correspondence*: for every Quinean ersatz world there is a genuine world with the represented pattern of occupancy and vacancy” what is more “the mathematical representations are a book-keeping device, to make sure the size and shape permitting proviso is satisfied.”<sup>87</sup> Clearly, though for Lewis there should not be any claim to equivalence between the representation and the concrete worlds that they represent.

In terms of concreteness one final aspect is raised by Lewis, this is the issue of alien individuals. Briefly, an alien property is a property that nothing in the actual world has but could have had none the less. When we consider all the possible individuals that a world might have had, it’s clear that some are parts of this world and others are not. We could have duplicates of any object with any property in any world but we needn’t necessarily have every individual instantiated in every world. When we consider complex wholes, constituted of parts of objects in worlds we may find that we have wholes constituted of parts, who are not duplicates of any part of this world. These parts are of course the ‘alien individuals’ and worlds which contain them are themselves alien individuals. Lewis calls these the ‘alien worlds’. We can look at all this in terms of properties. If a world has an object with a property that no object in the actual world has, and further that object can’t be built by combining objects in this world then that world has an alien property. It follows from this, that if something instantiates an alien property it is an alien individual and any world that has itself an alien individual is itself an alien world;



that is to say not our world. An example of this situation would be the various mythical beasts like unicorns and dragons. In the case of these mythical creatures, we see that they are only chimeras, that is, they are combinations of properties that exist in this world already; so a unicorn or a dragon can't really be an alien individual. Lewis argues that "it won't do to say that all worlds are generated by recombination from parts of this world, individuals which are possible because they are actual. We can't get the alien possibilities just by rearranging non-alien ones. Thus our principle of recombination falls short of capturing all the plenitude of possibilities."<sup>88</sup> Recombination then can't give us all the worlds required to fill the plenum of possible worlds, but that doesn't mean that the principle is totally inappropriate it still applies to all the worlds (both alien and non-alien) and is useful in generating the worlds we need for the possible world account. Of course if there are any alien properties at all then it should follow that there are many of them, and thus many alien worlds indeed.

### **Actuality**

We will now evaluate the last element of Lewis's account of the plurality of worlds, that of actuality. Lewis holds that the world we live in is in fact the actual while every other world is non actual but this should come as no surprise given that Lewis holds that actuality is matter of reference or point of view. What this means is that when someone from our world says that this world is actual, he certainly is being truthful but equally if someone from another world refers to his world as actual he too is being truthful. Lewis argues when I call this world actual "it applies to my world and my worldmates; to this world we are part of and to all parts of this world" what

is more if someone else uses it whether he be a worldmate of ours or whether he be unactualized” or not, then “it applies likewise to his world and his worldmates.”<sup>89</sup> This forms the basis of Lewis’s indexical account of actuality. What this means, is that saying this world is ‘actual’ is analogous to using words like ‘I’, ‘here’ and ‘now’. These terms gain their reference based on the circumstances they are uttered in and much like ‘actual’ they cannot fail to refer when they are used in the correct context. Lewis holds that “‘actual’ (in its primary senses) refers at any world  $w$  to the world  $w$ . ‘Actual’ is analogous to ‘present’ an indexical term whose reference varies depending on a different feature of context: ‘present’ refers at any time  $t$  to the time  $t$ .”<sup>90</sup> Actuality becomes then a relative matter. A world is actual but only to itself and its inhabitants and this applies to every single world equally. This shouldn’t surprise anyone given that Lewis’ account is one of the plurality of worlds, with every world on par with every other and no difference in kind between them.

If we entertain the assumption that one and only one world (presumably our own) is the only world which is actual and attempt to maintain a conception of the plurality of worlds, it becomes difficult to see how we could make sense of such a notion. Lewis raises two objections to this, the first regards our knowledge that we are actual and the second regards the contingency of such a statement. We will treat both here briefly.

Even if we can believe that the absolute actuality of this world stands for us, as Lewis notes, the relative distinction of other worldly individuals to their own actuality still exists. It would be a marvelous stroke of luck as Lewis points out, if of all the possible individuals, in all the possible worlds, throughout the pluriverse, it turned out that we happened to be part of the *real* actual world, as opposed to just believing that our world was actual when that was not the case. The inhabitants of other worlds, are in fact, just as we are in the so called actual world;

they love their mothers, pay their taxes and worry about their children, just as we do. So in those ways as with their own feeling of living in the actual world, they are just as we are. Clearly, we have a distinct notion that we are part of the actual world in the same way that we have a sense of the notion that this is indeed the world we are part of. Unfortunately, we can't make the leap from this to the notion that we are in fact the lucky inhabitants of the only truly actual world, there is no way we could ever know this. It follows from this that any person that could have been part of this world would have also had this feeling of living in the actual, thus in another possible world where this individual does exist it stands to reason that they would feel that the world they lived in was the actual one.

Concerning contingency, if we take actuality as an absolute matter it would be very hard not to hold that it was a contingent matter, why would this world need to be the actual world as opposed to another? Certainly, nothing forces this to be the case. Matters are contingent when they could have been otherwise, at one world something is the case at another it is not. It follows that at one world, a certain world is actual and at another, a different world might be actual, so it becomes difficult to maintain the absolute conception of actuality in light of this. Another issue that is of interest with regards to the contingency of actuality, concerns the question of whether the term 'actual' is rigid or not. That is to say should we consider the term 'actual' as having fixed reference or not, when we are discussing and considering other worlds. According to Lewis some indexical terms like 'now' are usually thought to be rigid in their designation where as others like 'present' seem to shift under certain circumstances. For him 'actual' is more like 'present' than 'now' in that sometimes it may be considered as rigid while at others it seems to be less rigid.

## Ersatz Modal Realism

We have considered the defining aspects of Lewis's version of Modal Realism and we saw that it commits us to a radically expanded account of ontology; where in, we find that for every way that the world could have been, there exist a world just like that somewhere in the multitude of worlds. What is more, each one of these worlds has full blown concrete existence on par with the actual world. Needless to say this is certainly an ambitious program that Lewis has committed himself to. It's worth noting that most of the support for the system of Modal Realism and its concrete worlds comes from its utility in various application in a multitude of fields in both science and logic. This line of argument is supposed to suggest that nothing so imminently useful could be false, however Lewis is quick to point out that if we are expecting to find the type of proof for their existence as full fledged concrete entities beyond their utility in a system we will be sorely disappointed. As Ian Hacking would say 'if you can spray them they are real', and in the case of the other worlds in Lewis's account we certainly cannot spray them. Does this mean that we should distance ourselves from the claim that the other worlds have concrete reality? Perhaps, but perhaps not, we might concede that we can have no direct contact or verification of their concrete reality,

while still holding that they are concrete as an article of faith, maintaining that the position does not imply any outright contradiction and it's merely unprovable in the final analysis. This leads to a type of deadlock between the two camps, those who can accept the concrete reality of the other worlds and those who cannot. The first group seems to found their position on the purposed utility of ascribing concrete existence to the other worlds, the second

group seems to disregard the possibility on the grounds that it is unparsimonious and leads to an overly crowded ontology. Of course there is a third possibility, and like most compromise position it leaves neither side very happy with the result. This is of course is the Ersatz account of possible worlds as abstract entities. Other worlds, do exist, but only as abstract constructs. The only world with truly concrete existence; will then be only the actual world.

We will now move to evaluate the Ersatz conception of possible worlds as abstract entities. Three distinct views can be put forward by the proponents of the Ersatz view, the first is *linguistic ersatzism*, the second is *pictorial ersatzism* and the third is *magical ersatzism*. We will have little to say about the second and third conceptions of Ersatzism beyond some broad comments, as they seem to be less credible than the first.

The Ersatz account is founded on the notion that instead of a multitude of concrete worlds we have but one concrete world and a vast array of abstract entities which can represent the way the world could have been. Accordingly, in the Ersatz conception, there is a firm line drawn between the concrete elements and the abstract elements that exist. We only have one concrete world which contains all the concrete entities that there are, other worlds only exist as abstract representations. In this way alternate possibilities are confined to the abstract realm where as Lewis puts it “we have license to believe in what we please.”<sup>91</sup> The Ersatzer claims that there is a multitude of abstract constructs which for him can play the same role as Lewis’s concrete entities. These ersatz worlds are abstractions which are meant to represent all the ways that the world could have been. Only one of these abstract representations properly represents the way the world is; in this way we can see that only one Ersatz representation of the world is *actualized* where as the rest are merely representations of the ways the world could have been and thus are *unactualized* corresponding naturally to the unactualized possibilities. This is of

course is a vast departure from Lewis's account, which holds that every single way that the world could have been is indeed a way that a world is. By this Lewis means to say that there is a concrete world for every possibility that exists. Not so for the Ersatz, the only thing which exists to correspond to an unactualized possibility is one more abstract representation among many. Along with whole unactualized worlds represented abstractly the Ersatz has also abstract unactualized individuals, these are of course the individuals that could have existed if things had been different and they are naturally the ones who populate the abstract ersatz worlds. As Lewis puts it "out of all the ersatz worlds and individuals there actually are, all but one of the worlds and the great majority of individuals are unactualized."<sup>92</sup> Of course it follows from this that according to itself each ersatz world is actual and all the others are merely the abstract representation. This feature of the Ersatz account can be seen as the relative actualization of the ersatz worlds. All the worlds in the Ersatz account, are equal in that each is actual relative to itself and none is actual relative to any other. This poses a problem for the Ersatz according to Lewis, for him the Ersatz "has not only relative actualization but absolute actualization besides so his ersatz worlds are not on equal footing after all. The one that represents the concrete world correctly isn't just actualized according to itself – it is actualized *simpliciter*."<sup>93</sup> The Ersatz might answer this by suggesting that there is only one ersatz world which is absolutely actualized. We aren't a part of that ersatz world because we are concrete and thus part of the concrete world, that ersatz world is an abstract representation and as such is merely an abstract facsimile of that concrete world.

The Ersatz worlds are there to replace the various possibilities in the analyses that we use. A possibility is said to be 'live' if and only if, there is some abstract ersatz world according to which it is found in. Lewis's Modal Realism can be rejected according to the Ersatz because

we have an equally compelling way of dealing with possibility that does not expand our ontology to proportions that common sense seem to forbid. Most Ersatzers however, would hold that they are not replacing possible worlds with abstract representation, but rather, possible worlds are in fact abstract representations. Lewis concedes that the Ersatz conception does have certain advantage because it is much more convivial to common sense in that it seems to be in agreement with what most people consider plausible in terms of what really exists and what does not. The ersatz representations can do the work in our analyses that the concrete worlds do, but Lewis contends that the price is high because the Ersatzer must resort to a primitive unanalyzable conception of modality where as the Modal Realist can offer a deconstruction of modality. Whether or not this is really the case, I leave as an open question for the reader to decide.

Lewis distinguishes three types of Ersatz Realism. The first is what he calls *Linguistic Ersatzism*, where the representations of the ersatz worlds are taken to be like stories or theories which are constructed from words of some language and they represent by way of their stipulated meanings. The second is what Lewis calls *Pictorial Ersatzism*, where the representations of the ersatz worlds are taken to be like pictures or diagrams, much like scale models and these represent by some sort of isomorphism. Third and finally there is *Magical Ersatzism*, where the representations do so simply by their nature and there is nothing more to say than that. Of these three the first seems to be the most plausible, the second seems implausible and the third even more so, to the point of being nearly incomprehensible. Given this we will now move to evaluate *Linguistic Ersatzism* leaving the other two conceptions of Ersatzism by the wayside.

Linguistic Ersatzism takes possible worlds as abstract entities, more precisely it takes the possible worlds to be maximally consistent sets of sentences which represent the various ways in

which the world could have been different. The ersatz worlds then represent by means of a world making language which in turn ‘says’ what the worlds are like, similar to a novel. If something is the case in a particular world; then it is so, if and only if, it is the case according to a novel or something like a novel. The ersatz worlds can be seen to be set-theoretic constructions made out of words. Each word is a set of spatiotemporal regions which give particular qualities and properties where they apply. In this way the ersatz worlds are set theoretic constructs made out of parts of the concrete world. As such the Linguistic Ersatz account is palatable because it provides an uncontroversial and commonsensical basis which most philosophers are already implicitly committed to, if not explicitly. Lewis concedes that this is indeed an advantage over his Modal Realism in that Linguistic Ersatzism concurs with opinions that we for the most part already hold.

To make Linguistic Ersatzim work for us, all that is required is a generalized language which would be a system of structures that will be used to interpret possible states of affairs in other worlds. As Lewis puts it “the words can be individuals that are part of the concrete world, or set-theoretic constructions out of those, or pure sets, or anything else we believe in.”<sup>94</sup> The account presented is closely related to that of Brian Skyrms who holds that the actual world is nothing more than the totality of facts, the same can be said for the other worlds. For Skyrms and others, the world then can be seen as nothing more than the set of these facts and facts are taken to be primitive irreducible entities, and these atomic facts are all we need to come to a conception of the worlds.

The mathematical universe provides an excellent way of characterizing the ersatz worlds. We can draw on the realm of pure sets to formulate our vocabulary in the world making language. In this way the program of Linguistic Ersatzism is in confluence with the previously



explored notion of taking possible worlds as mathematical representations in the spirit of Quine's Democritean worlds, as we saw earlier. We formulate the system as follows, we assume that space-time is flat, then we can assign each point a quadruple of real numbers. Additionally, we hold that a point in our representation can either be occupied or empty. Once we have a clear specification of the distribution of matter over all the points we will have a clear picture of each world, and thus no two worlds that are really different will be assigned the same representation. Concerning individuals in the ersatz worlds, a point should be stressed. A linguistic ersatz individual is strictly speaking not part of an ersatz world. Instead, an ersatz individual can be actualized according to an ersatz world. This is the case when the ersatz individual and the ersatz world are mirrors of one another.

### **Objections**

We will now move to evaluate some of the objection that Lewis raises with regards to Linguistic Ersatzism. The first of these objections is that the Linguistic Ersatzism takes modality as primitive and as such can offer no analyses of possible worlds. According to Lewis there are two ways that Linguistic Ersatzist must take modality as primitive. The first of these is that in the world making language of the Linguistic Ersatzist account not just any set of sentences will do, the type of set that will be required is a consistent one if not then it will fail to describe and represent any possible world. Lewis contends that consistency is a modal conception. He argues that when evaluating the consistency of a set of sentences we must judge which sentences *could* be held as true at the same time. This according to him is an irreducible modal conception. As

Lewis puts it “in order to say which things of the right nature – which sets of sentences of the world making language – are the ersatz worlds, we need to distinguish the consistent ones. That is *prima facie* a modal distinction: a set of sentences is consistent iff those sentences, as interpreted, *could* all be true together.”<sup>95</sup> Whether or not we should take this objection seriously is up for debate. It seems hard to agree with Lewis’s objection that consistency must be seen as a modal conception, when it seems to be obviously a syntactic one. Lewis also argues that implication must be taken to be a modal concept at its root. He argues that something might be the case according to a set of maximal consistent sentences not because it is explicitly stated as one of the sentences in the set but because it is implied by one or more of the sentences taken together. He argues that “this implication is *prima facie* modal: a set of sentences implies so-and-so iff those sentences, as interpreted, could not all be true together unless it were also true that so-and-so; in other words, if it is *necessary* that if those sentences are all true together, then so-and-so.”<sup>96</sup> Again this seems somewhat dubious, implication like consistency isn’t usually taken to be a modal concept.

Another issue in this vein of primitive modality, are the issues surrounding the axiom set of the ersatz worlds. The question is whether or not a proper axiom set can be specified without recourse to some notion of primitive modality. Lewis doubts that this is the case. There are two bad reasons for this according to Lewis both of which he had previously advanced. The first of these is the notion that the ersatz world should not be logically consistent yet mathematically inconsistent. So it follows from this that the axioms should include the mathematical truths at least as far as they are expressible in the language. As Lewis contends it “is impossible, of course, to specify them as exactly the sentences that meet some effective syntactic test; or even as exactly the theorems of some formal system. To specify them modally is to give up on the

plan of providing syntactic surrogates for primitive modality.”<sup>97</sup> The way around this according to Lewis is to take the mathematical axioms *model-theoretically*. The second of the bad reasons is the notion that it might not be possible to specify the axioms without recourse to modality itself and this engenders some type of vicious circularity. When choosing to add a truth to his set of axioms, the Ersatzer makes a decision about what statements are necessary and what statements are contingent. As the reasoning goes this is circular for the Ersatzer because it involves the very things he was attempting to define. This may seem convincing at first glance but as Lewis notes “circularity is a matter of what you analyze in terms of what; it is not a matter of why you think your analysis is right. It is not circularity if you build your analysis to give the answers it ought to, exercising your understanding of the *analysandum* as you go.”<sup>98</sup> It seems then that the axiom set can’t be had without being circular, at least not in this way.

So those are the bad reason for believing that we can’t specify the axiom set without recourse to primitive modality, so what of the good reasons? Lewis advances two of them. The first has to do with the fundamental properties and relations of simple things, the second has to do with the relation of global to local descriptions. We will examine each in turn. If we have a world making language that describes the world we will need two predicates for point particles one for ‘positive’ and one for ‘negative’ charges. The issue then revolves around the notion of consistency with regards to the ersatz worlds. Strictly speaking it will be consistent to say that a point particle can be both positively and negatively charged, because in purely syntactic terms all that is required for consistency is that in the language of the ersatz worlds, is that no atomic statement and its negation should be both included in the set of statements taken as true. One way around this according to Lewis is to specify an *axiom of unique charge*, which posits that nothing can have both positive and negative charge at the same time, and only ersatz worlds

which are in line with this axiom are consistent. However there is a problem with this axiom, it might be the case that the fact that no particle can have both charges is of course a law of nature, but in that case it is a purely contingent matter, it could have been different. This won't do because it will misrepresent modality, likewise its inclusion seems to smuggle primitive modality back into the system. For the second point regarding the relation of global to local descriptions we are faced with a much more serious objection than the previous one. When we set up the world making language to speak about spatiotemporal arrangements and fundamental properties of point particles we make the leap to assert features at the global level from these arrangements and fundamental properties at the local level. As Lewis puts it “we have a problem about consistency if the worldmaking language speaks both of local and global matters, both of particles and of donkeys. Or we have a problem about implicit representation if the worldmaking language speaks only of local matters, and yet ersatz worlds implicitly represent global matters.”<sup>99</sup> According to Lewis, what is required is *connecting axioms* which will relate the local properties of point particles to familiar global features, but this seems highly suspect as it implies such axioms can be constructed in finite ways; in fact, this is not the case, we seem not capable of specifying these axioms except by relying on some infinite set which will of course likely leave the Ersatz account of the axioms incomplete as it is impossible to construct .

With these two objections it seems that the Ersatz account must take the notion of modality as primitive. As Lewis points out most Ersatzers accept this consequence and willingly pay the price so to speak. For them, if the choice is between the radical ontology of Modal Realism and the primitive modality of Ersatz accounts, there is no contest and they will take an unanalyzed primitive modality. We will now move to briefly evaluate one final objection to the Linguistic Ersatz account, that of the descriptive power of the worldmaking language.

This objection can be divided into two parts. The first concerns the problem of indescrībables. Given that ersatz worlds and individuals are linguistic descriptions then it is never the case that we have two of them which are perfectly alike. By reducing ersatz possibility to its linguistic formulation in the worldmaking language then it follows that there can be only one for each linguistic formulation. It seems then, as Lewis suggests, that the ersatz world will then follow a simple principle of identity of indiscernibles with regards to possibilities. This will render the possibility in the Linguistic Ersatz account unsatisfactory as they will seemingly misrepresent different possible worlds as the same. We might easily dispense with this problem by stipulating against any indiscernible worlds in our Ersatz account, but the problem is not to be dealt with so easily, for individuals according to Lewis suffer from the same problem and we simply can't rid ourselves of this aspect of the problem by fiat. There are many situations where indiscernible individuals might emerge, however as Lewis puts it "we do not have correspondingly many indiscernible ersatz possible individuals, all actualized according to this ersatz world. One must do for all."<sup>100</sup> This is problematic because we can have many indiscernible possibilities, for an individual there is only one.

The second part of problem of the descriptive power of the worldmaking language concerns the need for the world making language to be specified by someone from the actual world; if this is the case then we can expect that it may fail to be sufficiently rich enough in terms of vocabulary to describe and distinguish all of the possibilities that will need to exist for the account to be complete. The reason for this, is that if we only have words for the natural properties from our world (the actual world), then it becomes doubtful that we will be able to describe the worlds that have alien properties not instantiated in the actual world. It's possible that our world is one that instantiates all properties that are possible and only other worlds are

ones lacking in properties but this seems unlikely and further we have no way of knowing either way. As Lewis puts it “there is no way the theorist can specify his world making language so that it will have words – names or predicates – for properties alien to his world. So any language he can specify will be inadequate to distinguish some of the possibilities there are.”<sup>101</sup> This might seem to be a decisive objection against the Linguistic Ersatz account, short of disallowing alien properties altogether the Ersatzer seems to have to accept some limitation on his system.

### **Ersatzism and Modal Realism on the Balance**

How should we evaluate the debate between the Ersatz account and Lewis’s Modal Realism? It seems we are stuck on the horns of a vicious dilemma. The Ersatz account seems to commit us to an unanalyzable primitive modality. On the other hand Lewis’s Modal Realism seems to commit us to an ontology which gives credence to a plethora of worlds which we seem unable to have any type of direct access to. This seems to be an argument concerning first principles. Do we accept the unanalyzable primitive modality or the expanded ontology? As James Felt has noted with regards to first principles in the debates concerning possible worlds, we can see that “there can be no knockdown arguments one way or another. Rather, we are at a level at which we must in effect play philosophic cops and robbers. “Bang, I got you!” “No, you missed me!” As the saying goes, you pays your money and you takes your choice.”<sup>102</sup> To return to Hacking for a moment:

“Now how does one alter the charge on the niobium ball? ‘Well at that stage’ said my friend, ‘we spray it with positrons to increase the charge or with electrons to decrease the

charge.’ From that day forth I’ve been a scientific realist. *So far as I’m concerned, if you can spray them then they are real.*” (Hacking p 22)

What is key here for us with regards to concrete possible worlds is the notion of *altering*, we can’t do this with the other possible worlds, the only world where we can alter something and then look for corresponding results is the actual one. Hacking argues accordingly:

“What convinced me of realism has nothing to do with quarks. It was the fact that by now there are standard emitters with which we can spray positrons and electrons – and that is precisely what we do with them. We understand the effects, we understand the causes, and we use these to find out something else. The same of course goes for all sorts of other tools of the trade, the devices for getting the current on the super cooled niobium ball and other almost endless manipulation of the ‘theoretical’.” (Hacking p 23)

If we cannot manipulate them or put them into cause and effect relations which can have observable effects we may need to admit that other concrete possible worlds will have to remain ‘empty’ theoretical entities.

Given this, when we are faced with this choice, how should we spend our intellectual capital? It depends on what matters to us, if we can accept entities that we can have no access to and live with philosophically clear consciences then we should do so. In that case the benefits will out way the costs as Lewis contends, we will have all the advantages of the Modal Realist account in terms of utility and we need only take on the other worlds as concrete entities in our ontology. If on the other hand we feel that the price is too high so to speak, that the claim that other worlds are concrete is simply too hard to swallow in terms of ontology then the only avenue open to us to retain a workable account of modality will be the Ersatz account. It seems a

hard choice to make between these two positions, however something is worth underlying. The Modal Realist account forces us to accept an unobservable and unverifiable entity, in fact not just one entity but infinite numbers of these entities. If we accept one kind of unobservable and unverifiable entity, why not another? If another theory of science claims to have explained a vast domain of intellectual pursuit by recourse to an unobservable and unverifiable entity how will we rule on that entity? If the entity in question is like the other worlds of Lewis's account in that they do not imply an explicit contradiction and lead to useful results in terms of explanation, can they thus be appropriated into the body of scientific knowledge as reasonable accounts? This is somewhat troubling because we will be forced to admit that we have no better or worse reasons to reject these new entities, indeed, if we allow for other concrete worlds beyond the reach of observation and verification other entities like this won't be far behind.

We can't pick and chose, either we accept these type of unobservable and unverifiable entities as respectable scientific candidates for our accounts of phenomenon or we don't. Clearly we can't differentiate between various unobservables as better or worse than each other for our accounts, they will all be on equal footing, the only measure will then be their utility in explanation. On this point alone the account of concrete alternative worlds seems to entail something that common sense cannot accept. If it were the case that the Modal Realist's account only left the door open to their concrete worlds without leaving the possibility of other unobservables and unverifiables this might be acceptable, after all, what is the harm of one such entity, but this isn't the case. Given this we might have to make do with the Ersatz account, it may be less glamorous and even more seriously, less useful than Lewis's account but clearly it puts us on more secure and safe methodological grounds and for that alone it seems deserving of our allegiance.



## Conclusion

We have now completed our study of these issues. First, we explored the distinction between indicative and counterfactual statements and more broadly the nature of conditionals. Second, we gave a detailed account of David Lewis's system of counterfactuals. Finally, we evaluated the metaphysical underpinnings of the possible worlds account paying special attention to its relationship to Lewis's system of counterfactuals. We will now briefly summarize the issues presented in each section.

In the first section we looked at the distinction between indicative and subjunctive statements for conditionals. The distinction was fraught with problems and certain authors have suggested that we collapse the distinction. They argued that issues with regards to the temporality of the sentences themselves make it difficult to uphold any kind of distinction between the two. With regards to counterfactual statements, which can be seen to be a sub species of subjunctive statements, we saw that the issue was first raised by the Logical Positivists who hoped to more adequately found the notion of scientific law on an account of counterfactuals. For these Logical Positivists the problem centered around the familiar dispositional properties like fragility and solubility. With regards to the distinction we saw that despite the problems associated it still warrants our allegiance, at least for the time being. We also evaluated the issue of whether or not conditionals can be taken to be propositions. If they are indeed propositions then they conceivably can have truth values like all other propositions and this truth value will be one that is founded on the truth functional account. This is to say its

truth value will be the result of the value of its constituent parts, in the case of the conditional, the antecedent and consequent. I argued that the conditional is not a proposition. Different attempts were made to secure the truth functional account of conditionals with varying results. Finally we briefly evaluated the issues surrounding conditionals and fact stating discourse. It was found that no clear consensus has been reached on this issue despite extensive study by various authors. I suggested that perhaps we can say that conditionals despite not being propositions could still be part of fact stating discourse, in the sense that they are used projectively as epistemic strategies projected on the world.

In the second section we looked at David Lewis's system of counterfactuals. Lewis introduced two counterfactual operators and argued that we can establish the truth conditions for sentences with these operators by recourse to an account of possible worlds where in the antecedent and consequent will be given truth values at various possible worlds and thus the counterfactual can be evaluated in terms of these values. Lewis argued that the conditional in counterfactuals can be seen to be a kind of variably strict conditional. We also evaluated the centering conditions of Lewis's system of counterfactuals and looked at the subsequent truth conditions for counterfactuals that Lewis assigned based on these. We looked at the four cases for truth and falsity of counterfactuals. We saw that there were two which give a value of true and two which give the value of falsity. The first of the two cases are vacuous truth and non vacuous truth and the second two cases are falsity, opposite true and falsity opposite false. We also looked at Lewis' treatment of the important principle of the Limit Assumption in his system. This led to amended truth conditions which respect the Limit Assumption's stipulations. We also evaluated Lewis's attempt at giving more precise interpretation to the modal aspects of the system and also looked at the issue of impossible antecedents. We also evaluated the notions of

inner and outer modalities in the centered system of spheres. Finally we looked at the counterfactual fallacies and some criticisms of Lewis's account of them.

In the third section we looked at Lewis's account of Modal Realism and the rival account of Modal Ersatzism. We explored some of the issues raised by Lewis such as isolation, concreteness, plenitude and actuality. With regards to isolation we saw that each world of the pluriverse is causally isolated from one another. With regards to concreteness we saw that each of the other worlds of the pluriverse is thought to have as much concrete reality as the actual world. With regards to plenitude we saw that the account of Modal Realism allows for a vast plenitude of possible worlds through Lewis' principle of recombination. With regards to actuality we saw that every single world of the pluriverse is actual in relation to its self and there are no absolutely actual worlds. Finally we evaluated the Modal Ersatz account which holds that the other possible worlds are only abstract constructs and not concrete totalities as the Modal Realist account holds. We compared the two accounts and found that for reasons of parsimony the Modal Realist account is untenable as it makes use of entities that are both unobservable and unverifiable. Perhaps we may need to admit what old Duhem knew all along; not everything which is consistent is deserving of our allegiance.

---

---

## Citations

- <sup>1</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 236
- <sup>2</sup> VH Dudman, Indicative and Subjunctive, *Analysis*, Vol 38 No. 3, 1988 p 122
- <sup>3</sup> Ernest Adams, Subjunctive and Indicative Conditionals, *Foundations of Language*, Vol. 6 No. 1 1970 p 93
- <sup>4</sup> Kit Fine, Critical Notice, *Mind New Series*, Vol. 84 No. 335, 1975, p 451
- <sup>5</sup> Kit Fine, Critical Notice, *Mind New Series*, Vol. 84 No. 335, 1975, p 451
- <sup>6</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 240
- <sup>7</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 239
- <sup>8</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 242
- <sup>9</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 245
- <sup>10</sup> James F. Thomson, In Defense of  $\supset$ , *The Journal of Philosophy*, Vol. 87 No. 2, 1990, p 57
- <sup>11</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 247
- <sup>12</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 281
- <sup>13</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 280
- <sup>14</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 284
- <sup>15</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 300
- <sup>16</sup> DH Mellor, How to Believe a Conditional, *Journal of Philosophy*, Vol. 90 No. 5, 1993, p 236
- <sup>17</sup> Robert Stalnaker, A Theory of Conditionals, *Studies in Logical Theory*, (N. Rescher, ed.), Oxford, 1968, p 103
- <sup>18</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 310
- <sup>19</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 311
- <sup>20</sup> Sabine Iatridou, The Grammatical Ingredients of Counterfactuality, *Linguistic Inquiry*, Vol. 31 No. 2, 2000, p 231
- <sup>21</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 313
- <sup>22</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 313
- <sup>23</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 313
- <sup>24</sup> Dorothy Edgington, On Conditionals, *Mind*, Vol. 104 No. 414,1995, p 319
- <sup>25</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 1
- <sup>26</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 3
- <sup>27</sup> Kit Fine, Critical Notice, *Mind New Series*, Vol. 84 No. 335, 1975, p 453
- <sup>28</sup> Kit Fine, Critical Notice, *Mind New Series*, Vol. 84 No. 335, 1975, p 453
- <sup>29</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 8
- <sup>30</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 9
- <sup>31</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 9
- <sup>32</sup> Kit Fine, Critical Notice, *Mind New Series*, Vol. 84 No. 335, 1975, p 455
- <sup>33</sup> Kit Fine, Critical Notice, *Mind New Series*, Vol. 84 No. 335, 1975, p 456
- <sup>34</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 10
- <sup>35</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 10
- <sup>36</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 11
- <sup>37</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 12
- <sup>38</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 13
- <sup>39</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 13
- <sup>40</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 14
- <sup>41</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 14
- <sup>42</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 15
- <sup>43</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 16
- <sup>44</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 16
- <sup>45</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 19
- <sup>46</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 20
- <sup>47</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 21
- <sup>48</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 21
- <sup>49</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 23
- <sup>50</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 24
- <sup>51</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 26

- 
- <sup>52</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 27
- <sup>53</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 28
- <sup>54</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 29
- <sup>55</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 29
- <sup>56</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 30
- <sup>57</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 30
- <sup>58</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 31
- <sup>59</sup> Jeffrey Dunn Fried Eggs Thermodynamics and the Special Sciences, *BJPS* 2006, p 2-3
- <sup>60</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 33
- <sup>61</sup> EJ Lowe, *A Survey of Metaphysics*, Oxford, 2002, p 144
- <sup>62</sup> EJ Lowe, *A Survey of Metaphysics*, Oxford, 2002, p 151
- <sup>63</sup> EJ Lowe, *A Survey of Metaphysics*, Oxford, 2002, p 152
- <sup>64</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 2
- <sup>65</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 87
- <sup>66</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 2
- <sup>67</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 3
- <sup>68</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 4
- <sup>69</sup> Allen Stairs, Review Essay ‘ On the Plurality of Worlds’, *Philosophy and Phenomenological Research*, Vol. 49 No. 2, 1988, p 339
- <sup>70</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 4
- <sup>71</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 70
- <sup>72</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 71
- <sup>73</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 76
- <sup>74</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 78
- <sup>75</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 80
- <sup>76</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 81
- <sup>77</sup> Allen Stairs, Review Essay ‘ On the Plurality of Worlds’, *Philosophy and Phenomenological Research*, Vol. 49 No. 2, 1988, p 335
- <sup>78</sup> Brian Skyrms, Possible Worlds, Physics and Metaphysics, *Philosophical Studies*, Vol. 30 No. 5, 1976, p 326
- <sup>79</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 81
- <sup>80</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 81
- <sup>81</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 84
- <sup>82</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 86
- <sup>83</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 88
- <sup>84</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 89
- <sup>85</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 90
- <sup>86</sup> David Lewis, *Counterfactuals*, Blackwell, 1973, p 90
- <sup>87</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 91
- <sup>88</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 92
- <sup>89</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 92
- <sup>90</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 93
- <sup>91</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 137
- <sup>92</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 138
- <sup>93</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 138
- <sup>94</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 144
- <sup>95</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 151
- <sup>96</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 151
- <sup>97</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 153
- <sup>98</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 154
- <sup>99</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 155
- <sup>100</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 157
- <sup>101</sup> David Lewis, *On the Plurality of Worlds*, Blackwell, 1986, p 159
- <sup>102</sup> James Felt, Why Possible Worlds Aren’t, *The Review of Metaphysics*, Vol. 50 No. 1, 1996, p 68

---

## Bibliography

Adams, E. W. 1965: "A Logic of Conditionals". *Inquiry*, 8 pp. 166-197.

Adams, E. W. 1970: "Subjunctive and Indicative Conditionals". *Foundations of Language*, 6 pp. 89-94.

Anderson, Alan Ross 1951: "A Note on Subjunctive and Counterfactual Conditionals". *Analysis*, 12 pp. 35-38.

Arif, Ahmed 2007: *Saul Kripke, Contemporary American Thinkers*, Continuum.

Ayers, M. R. 1965: "Counterfactuals and Subjunctive Conditionals". *Mind* 7, pp. 347-364.

Barker, John, A 1969: *A Formal Analysis of Conditionals*, Southern Illinois University Monographs, Humanities Series, 3.

Bennett, Jonathan 2003: *Philosophical Conditionals*, Oxford University Press

Bennett, Jonathan 1974: Review of David Lewis, *Counterfactuals*. *Canadian Journal of Philosophy*, 4, pp. 381-402.

Carnap, Rudolf 1932: The Elimination of Metaphysics Through the Logical Analysis of Language, *Erkenntnis*, 2.

Carnap, Rudolf 1947: *Meaning and Necessity*, University of Chicago Press.

---

Cartwright, Nancy 1999: *The Dappled World, A Study of the Boundaries of Science*, Cambridge University Press.

Chellas, Brian 1980: *Modal Logic: An Introduction*. Cambridge University Press.

Chisholm, R. M. 1946: "The Contrary-to-fact Conditional". *Mind*, 55, pp. 289-307.

Copeland, Jack 2006: "Meredith Prior, and the History of Possible World Semantics" *Synthesis*, 3, 150 pp. 373-397.

Dudman, V. H. 1983: "Tenses and Time in English Verb Clusters of the Primary Pattern". *Australian Journal of Linguistics*, 3, pp. 25-44.

Dudman, V. H. 1988: "Indicative and Subjunctive". *Analysis*, 48, pp.113-122.

Dudman, V. H. 1994: "On Conditionals". *Journal of Philosophy*, 91, pp. 113-128.

Dunn, Jeffery 2010: Fried Eggs, Thermodynamics and the Special Sciences. *British Journal of the Philosophy of Science*. pp 1-38

Edgington, Dorothy 1986: "Do Conditionals Have Truth Conditions?", *Crítica: Revista Hispanoamericana de Filosofía*, Vol. 18, No. 52 pp 3-39.

Edgington, Dorothy 1995: "On Conditionals". *Mind*, 104, pp. 235-329.

Ellis, Brian 1984: "Two Theories of Indicative Conditionals". *Australasian Journal of Philosophy*, 62, pp 50-66.

---

Ellis, Brian 2001: *Scientific Essentialism*, Cambridge University Press.

Felt, James 1996: Why Possible Worlds Aren't, *The Review of Metaphysics*, Vol. 50 No. 1, , pp. 63-77.

Fine, K. 1975: Critical notice of David Lewis's *Counterfactuals*. *Mind*, 84, pp. 451-458.

Gibbard, A. 1981: "Two Recent Theories of Conditionals" in Harper, Stalnaker and Pearce (eds.) 1981, pp 211-247.

Goodman, N. 1947: "The Problem of Counterfactual Conditionals". *Journal of Philosophy*, 44, pp. 113-128.

Hacking, Ian, 1986: *Representing and Intervening. Introductory Topics in the Philosophy of Natural Science*. Cambridge University Press.

Iatridou, Sabine, 2000: The Grammatical Ingredients of Counterfactuality, *Linguistic Inquiry*, Vol. 31 No. 2, pp. 231-270.

Jackson, Frank 1987: *Conditionals*. Oxford: Basil Blackwell.

Jefferey, Richard, Edgington, Dorothy, 1991: Matter-of-Fact Conditionals, *Proceedings of the Aristotelian Society, Supplementary Volumes*, Vol. 65. pp.161-183 and pp.185-2009

Kripke, Saul 1963: "Semantical Considerations on Modal Logic." *Acta Philosophica Fennica* 16, pp. 83-94.

Kripke, Saul 1972: *Naming and Necessity*, Blackwell.



Leibniz, G.W 1989: *Philosophical Essays*, translated by Roger Ariew and Daniel Garber, Hackett Publishing Company.

Lewis, David, 1973: *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, David 1979: 'A Problem about Permission', in *Essays in Honour of Jaakko Hintikka*, ed. by E. Saarinen et al., 163-175

Lewis, David, 1986: *On the Plurality of Worlds*, Blackwell.

Lewis, David 1986: *Philosophical Papers*, volume II. Oxford.

Lowe, E.J. 1990: "Conditionals, Context and Transitivity". *Analysis*, 50, 2, pp. 80-87

Lowe, E.J. 1995: "The Truth about Counterfactuals". *Philosophical Quarterly*, 43, pp 41-59.

Lowe, E.J. 2002: *A Survey of Metaphysics*, Oxford Press.

Martin, Robert 1982: *The Meaning of Language*, MIT Press 1982.

Mellor, D.H. 1993: "How to Believe a Conditional". *Journal of Philosophy*, 90, 5, pp. 233-248.

Skyrms, Brian 1976: "Possible Worlds, Physics and Metaphysics", *Philosophical Studies*, Vol. 30 No. 5, pp. 323-332.

---

Stairs, Allen 1988: Review Essay 'On the Plurality of Worlds', *Philosophy and Phenomenological Research*, Vol. 49 No. 2, pp. 333-352.

Stalnaker, R. 1968: "A Theory of Conditionals" in *Studies in Logical Theory, American Philosophical Quarterly*, Monograph Series, 2. Oxford: Blackwell, pp. 98-112.

Stalnaker, R. 1970: "Probability and Conditionals" *Philosophy of Science*, 37, 1, 64-80.

Strawson, P.F. 1986: "'If' and '□'" in Gandy, R. E. and Warner, R. 1986, pp. 229-242.

Thomason, James 1990: "in Defense of □". *Journal of Philosophy*, 87, pp. 56-70.

Tichy, Pavel 1976: "A Counterexample to the Stalnaker-Lewis Analysis of Counterfactuals" *Philosophical Studies*, 29, pp. 271-273

Woodward, James 2005, *Making Things Happen, A Theory of Causal Explanation*, Oxford University Press.

Wright, Crispin 1983: "Keeping Track of Nozick". *Analysis*, 43, 3, pp. 134-40