THE METAPHYSICS OF CAUSATION

CAUSATION, CONDITIONALS AND REGULARITIES

by

Peter Richard Haddow Forrest, B.A. (Oxon.),
Ph.D. (Harvard).

submitted in fulfilment of the
requirements for the degree of

Master of Arts

UNIVERSITY OF TASMANIA

HOBART

OCTOBER 1978

TABLE OF CONTENTS

* * * * * * *

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university, and, to the best of my knowledge and belief, this thesis contains no copy or paraphrase of material previously published or written by another person, except where due reference is made in the text of the thesis.

Signed: P.R.A. Faust.

Date : October 2nd. 1978

ACKNOWLEDGEMENTS

I am grateful to the following persons who have helped me with this thesis: Professor W.D. Joske and Dr. E.E. Sleinis of the University of Tasmania, Dr. M. Tooley of the Research School of Social Sciences at the Australian National University, Mrs. N.M. Gill the Philosophy Department secretary at the University of Tasmania, and my wife Felicity.

Professor Joske supervised my research on the nature of Metaphysics from which this thesis is derived and he has made several helpful comments on the thesis itself. Dr. Sleinis who supervised the thesis itself, has patiently and thoroughly commented on this work and has consequently enabled me to remove many obscurities. Dr. Tooley kindly read a rough draft of the bulk of the thesis and made some very useful criticisms. Mrs. Gill typed a draft and the final version of the thesis which necessitated working from an untidy and at times almost illegible manuscript. Lastly, Felicity has not only supported and encouraged me in my work but has also typed a draft of the thesis and advised me in matters of punctuation, orthography and style. The remaining stylistic infelicities are entirely due to my stubbornness.

A B S T R A C T

In this thesis I am concerned not with *immanent* causation, but with *transeunt* causation.

In the Introduction I show how the *Conditional Analysis* and the *Regularity Account* of causation can be used to defend *Causal Anti-realism* and to supplement the *Manipulability Account*. I also mention three theses about Metaphysics which my discussion of causation illustrates.

In Chapter One, I discuss some preliminaries. In Section One, I consider the method to be used for testing the *permissibility* of accounts of causation. I provide a *stricter* and a *less strict* criterion: the *Analysis-criterion* and the *Reduction-criterion*. The former is used in Chapter Two, the latter is used in Chapter Three. In Section Two, I argue for a distinction between two kinds of transeunt causation : *explanatory* and *producing* causation. I concentrate on producing causation in this thesis. I also argue that the class of particulars which can be *producing causes* - I call this the class of *quasi-events* - is larger than the class of events. Quasi-events are defined in terms of the notion of a *causal field*, which I discuss and defend. In Section Three, I briefly expound the kinds of conditional required for the Conditional Analysis.

Chapter Two has two parts. In Part One, I defend the Conditional Analysis of causation. I begin with Mackie's account in *The Cement of the Universe* and I successively modify this account to meet various objections. The objections considered include the problems raised by *collateral effects*, *overdetermination* and

simultaneous causation. I also briefly consider Lewis' account of causation. In Part Two, I discuss what account might be given of the conditionals used in the Conditional Analysis. I conclude that these conditionals should be treated as *meta-inference* conditionals based on a special kind of *non-deductive* inference.

Chapter Three also has two parts. In Part One, I defend the possibility of a regularity account, provided it is to satisfy only the Reduction-criterion. In Part Two, I provide a regularity account based on Mackie's account of causal regularities in *The Cement of the Universe*. Objections similar to those of Chapter Two are discussed.

Finally, in Chapter Four, I discuss backwards causation, which I ignore in Chapters Two and Three. I argue that a causal anti-realist can quite rationally hold the position that backwards causation is logically possible but metaphysically impossible. More precisely, there are no coherent examples, described in non-causal terms, which a causal anti-realist need interpret as cases of backwards causation.

INTRODUCTION

In this thesis I shall be concerned not with immanent causation, in which an agent directly causes an effect, but with transeunt causation, in which one event or state-of-affairs causes another event of state-of-affairs. I have two reasons for discussing transeunt causation. *My first reason* is that I wish to defend a regularity account of causation in the tradition of Hume[1] - I also defend another, compatible, account of causation in terms of conditionals, in the tradition of one of Hume's definitions of cause:

> ... we may define a cause to be an object followed, by another ... where, if the first object had not been, the second never had existed.[2]

My second reason is that my discussion of causation illustrates, and perhaps to some extent confirms, three theses about Metaphysics and its relation to ordinary language.

0.1 First I shall discuss why I consider it important to give a conditional analysis and a regularity account of transeunt causation. These two accounts of causation are used to defend the anti-realist view of causation. Causal Anti-realism, expressed in the language of Hume, is that an account of causation "in the objects" can be given in non-causal terms and that in any sense in which the cause necessitates the effect, causal necessity is "in the mind". Another way of stating the anti-realist view of causation would be to claim that causation is not a primary but a secondary relation between events. Of course, the causal anti-realist is not denying that one event can

---

1. Hume, D., *A Treatise of Human Nature*, Book 1, Section 14.
2. Hume, D., *Enquiries Concerning the Human Understanding*, Section VII, Pt.2.

cause another, nor is he denying that there is causation "in the objects". The causal anti-realist is merely denying that there is any irreducible element "in the objects" such as a mysterious causal necessity or power. Even leaving aside any difficulties with the notions "in the objects" and "in the mind", the anti-realist view is as vague as the phrase 'non-causal term' is. Are, for example, nomic necessity and physical probability causal notions? If they are, then someone giving an account of causation in terms of nomic necessity is a causal realist. This vagueness is of no great consequence provided one supplements the anti-realist view by a positive account of causation "in the objects". In Part 2 of Chapter Two on the Conditional Analysis, I shall propose an account of causation in terms of inferences and non-accidental generalisations (non-accidental generalisation might be considered a causal notion); in Chapter Three I propose a regularity account of causation in terms of indisputably non-causal notions.

I shall now examine two initially plausible accounts of causation which might *seem* to make the Conditional Analysis and the Regularity Account at best unnecessary labour spent in defence of Causal Anti-realism, and at worst either incorrect or misguided.

0.2 The Singularist Account

The singularist adopts the straightforward policy of carefully examining individual causal situations and, he claims, finding necessary and sufficient conditions for event C to cause event K. Ducasse has given an account of this kind:

> Considering two changes, C and K ..., the change C is said to have been sufficient to, i.e. to have caused, the change K, if:
> 1. The change C occurred during a

time and through a space terminating at the instant I at the surface S.

2. The change K occurred during a time and through a space beginning at the instant I at the surface S.
3. No change other than C occurred during the time and through the space of C, and no change other than K during the time and through the space of K. [3]

If C and K satisfy these three conditions I shall say that they satisfy *the Ducasse criterion*. Satisfying the Ducasse criterion certainly seems to be a good evidence that C causes K. Indeed, one might even accept that satisfying the Ducasse criterion is a *part* of a sufficient truth-condition for C to cause K. More precisely, if the Ducasse criterion is satisfied but C does *not* cause K, then there must be some kind of outside interference. Now if the Ducasse criterion is satisfied, this interference must *not* be considered to result in a (macroscopic) change occurring through the space of process C or process K; so the interference cannot occur near S. *Either* the interference involves the direct action of a spirit (immanent causation presumably) *or* it involves action-at-a-distance, *or* it involves a (non-causal) process deemed not to be considered a change for the purposes of condition 3. Action-at-a-distance is ruled out by the Ducasse criterion, so this leaves as possible kinds of interference the intervention of a spirit and non-causal processes. Suppose, when giving the account of causation, only perceptible macroscopic processes are considered as changes. Then processes which are explained in terms of the theoretical entities of Physics could be used to interfere with the situation; these would be examples of interfering non-causal processes. But, I shall now argue, the Ducasse criterion cannot provide necessary and sufficient truth-conditions for causation and hence cannot

---

3. Ducasse, C.J., *Truth, Knowledge and Causation*, pp. 3, 4.

provide a list of features on which causation is supervenient. For suppose a hollow metal sphere is at rest, apparently unsupported, three feet above the ground, and then it begins to move sideways. Suppose that there is no interference, say by means of magnets. Suppose also that the only change to occur in the vicinity of the sphere is that a significant proportion of the electrons in the sphere change their spin. The advocate of the Ducasse criterion as a necessary condition is now faced with a dilemma: does he take into consideration the theoretical entities of Physics or not, when using the Ducasse criterion? I shall now consider four replies:

Reply One: A operationalist might reply that the theoretical entities of Physics do not exist so the movement of the sphere has no cause. Thus the Ducasse criterion is vindicated. But now consider the case in which electromagnets *are* used to move *magnetic* spheres. According to the operationalist, the cause of the motion of the magnetic spheres is the switching on and off of the magnets. Thus if there are no theoretical entities of Physics, these must be cases of action-at-a-distance. So the operationalist is committed to action-at-a-distance; hence the defender of the Ducasse account is committed to the denial of Operationalism.

Reply Two: One might reply that submicroscopic processes are not considered to be changes, although the theoretical entities of Physics do exist. In this case, a non-causal process results in the motion of the sphere. But there is a difficulty with this reply. Suppose that, as is conceivable, the fundamental particles were very small hard spheres obeying the laws of Newtonian Mechanics. Then one particle could hit

another so that the first stops moving and the second starts without, according to this reply, there being any causation. It would seem peculiar if the notion of causation were relevant to the macroscopic situation but not to a submicroscopic situation identical in all respects but scale.

Reply Three: The theoretical entities of Physics are always taken into account when using the Ducasse criterion. In this case the Ducasse criterion never results in simply described causes and effects. For there are always many processes occupying the same region as C; either there will be changes in fields or there will be processes involving the movement of subatomic particles.

Reply Four: The events occurring in or to the theoretical entities of Physics are initially considered not to be changes and to provide *non-causal* processes. But if (using the Ducasse criterion) no cause is found, *then* one considers these theoretical entities. In this case, when there is no macroscopic cause, there would be too many candidates for the cause under consideration, and hence the true cause on the Ducasse criterion would involve excessive and unnecessary detail. Thus perhaps the laws of Nature entail that the change in electron-spin is followed by the movement of the sphere, yet a quite irrelevant change in the spin of neutrons which also occurred would also have to be considered part of the cause.

Although the Ducasse criterion is important as evidence, it fails as an account of causation: I suspect that any singularity theory of causation will also fail to provide necessary and sufficient

truth-conditions for causation and hence fail to provide a list of non-causal features on which causation is supervenient.

0.3 The Manipulability Account

One way of defending Causal Anti-realism is based on the Manipulability Account of Causation, in which event C causes event E if by bringing C about one can bring E about, or if by preventing C one can prevent E.[4] One might give the following version of this defence of Causal Anti-realism.

Hume's regularity account is based on a misguided response to the discovery that causation is a supervenient (consequential) relation: the world could not be altered so that A did not cause B without altering some other features of the world. An account of causation "in the objects" is an account of those features on which causation is supervenient. Hume's mistake was to attempt the task of explicitly listing those features. Any attempt to make such a list is either unsuccessful because there are counter-examples, or unsatisfactory because it becomes excessively complicated. (My version of the Regularity Account could be considered an example of this excessive - one might even say tedious - complexity.) The hard work of explicitly describing, without using anthropocentric terms, the features on which causation is supervenient is unnecessary for two reasons.

The first reason is that the supervenience of causation is itself good evidence against causal realism. One can be confident that one could *in principle* provide an adequate account of the features on which causation is supervenient; bearing in mind that the

---

4. See Gasking, D., "Causation and Recipes", *Mind*, Vol.64, (October 1955), pp. 479-487.

term 'cause' might have some degree of open-texture.

The second reason is that an adequate, simple and informative account of causation can be given by considering people's purposes in using causal language. Compare causation with tablehood. Tablehood is supervenient on the (non-tabular) properties of a material object. Here one might make the proviso that a table be manufactured or at least, if it is a naturally occurring object, selected. Also, I assume, for the sake of giving an example, that it is not true by definition that a table is *used as a table;* a table could be used as a work of art and still be a table, and not merely *like* a table. One might find an explicit list of the features on which tablehood is supervenient; but this is a misguided approach, for one can discuss these features without actually listing them. Thus one might begin with the following account of tablehood.

Account of Tablehood, One:

In addition to being either manufactured or selected, a table is a material object which has features $F_1 \ldots F_m$ (whatever they may be) which make it eminently suited for use-as-a-table. A material object M is used-as-a-table if one or more people sit near M and put down or pick up small objects on M (small enough to be manipulated easily with one hand), without undue stretching or bending, but they do not put objects in containers inside M. (I hope the reader will ignore any inadequacies in this account of 'use-as-a-table'.)

It might be objected that one could have a table ten feet high. It is debatable whether such a giant's table is a table, or whether it is merely something like a table. Nonetheless, one might modify one's account of tablehood to cover all such counter-examples in

either of two ways. One would then have:

Account of Tablehood, Two:

A material object is a table if it is manufactured or selected and if it has *most* of the features $F_1 \ldots F_m$ (whatever they may be) such that a material object with *all* the features of $F_1 \ldots F_m$ is eminently suited for use-as-a-table.

Account of Tablehood, Three:

The *paradigms* of tables are manufactured material objects which are eminently suited for use-as-tables. A material object is a table if it is manufactured or selected and if it is sufficiently *similar to the paradigms*.

Surely some such account is as satisfactory as one could require of the features $F_1 \ldots F_m$ on which tablehood is supervenient.

Causation is related to the activities of manipulation, explanation, and prediction, as tablehood is related to the activity of using a material object as-a-table. Thus one obtains an initially extremely attractive account of the features on which causation is supervenient:

Account of Causation, One:

In addition to the requirements that A and B occur, that A and B are distinct existences and (perhaps) that A is not later than B, A causes B if and only if features $F_1 \ldots F_m$ (whatever they may be) are present in the situation which ensure that:

(1) Bringing A about could be used as a means of bringing B about, and/or preventing A could be used as a means of preventing B; *and/or*

(2) The occurrence of B can be explained in terms of the occurrence of A; *and/or*

(3) The occurrence of A can be used to predict the occurrence of B.

This account needs some modification. For if A and B are collateral effects of some third event C and A occurs before B, then the occurrence of A can be used to predict the occurrence of B. For example, the flowering of one species of tree in the rain-forest might be used to predict the ripening of the fruit of another species a few days later; both being effects of an imperceptible change in temperature. Therefore sufficient condition (3) should be discarded. Also, as I shall argue in Chapter One, Section Two, there are two distinct notions of transeunt causation: *explanatory causation* in which one fact explains another fact; and *producing causation* in which, in the paradigms, one event causes another event. In this thesis I shall proceed as if the notion of explanation is unproblematic in order to avoid a lengthy discussion of explanatory causation; I shall concentrate on producing causation. Thus one obtains the following account of producing causation:

Account of Causation, Two:

In addition to the requirements that A and B are distinct existences, that A and B occur and (perhaps) that A is no later than B, A causes B if and only if features $F_1 \ldots F_m$ (whatever they may be) are present which ensure that bringing A about could be used as a means of bringing B about, or preventing A could be used as a means of preventing B.

Account Two needs further modifications for there are two objections to it as it stands:

Objection One: In some situations it is technically impossible to bring about A, so how can one bring about A in order to bring

about B? For example, one might say that the explosion of a star caused the presence of a nebula. To this it might be replied that in principle exploding a star could be used to produce a nebula or that, if God exists, God could explode a star in order to produce a nebula. This reply is unsatisfactory, for it seems that the grounds one has for saying that in principle one could explode a star in order to produce a nebula, or for saying that God could explode a star in order to produce a nebula, is that exploding the star *causes* the nebula. Hence someone might argue that the features which ensure that in principle one could explode a star in order to produce a nebula include the generalisation that explosions of stars *cause* nebulae. So the account is circular.

Objection Two: There is an important example due to Chisholm[5] in which someone raises his arm in order to make certain events occur in his arm a little earlier. Raising the arm is here considered to be a movement of the arm, not the intention to move the arm. Suppose that these neural events consist of small electric currents which can be made to sound an alarm, say by using a radio-transmitter built into the arm. Suppose the alarm sounds later than the raising of the arm. Then one could raise one's arm as a means to sounding the alarm, and one could refrain from raising one's arm in order to prevent the alarm sounding. Yet surely the raising of the arm does not cause the alarm to sound; both are *collateral effects* of the neural events. Hence there seem to be situations which, according to Account Two, should be cases of cause and effect but are,

---

5. Chisholm, R.M., "Freedom and Action", in *Freedom and Determinism*, ed. Lehrer, K., pp. 43, 44.

in fact, cases of collateral effects.

To meet these three objections I modify Account Two to obtain:

Account of (Producing) Causation, Three:

In addition to the requirement that A and B occur, that A and B are distinct existences and (perhaps) that A is no later than B, A causes B if and only if there are present features $F_1 \ldots F_m$ in virtue of which the situation is *similar* to the *paradigms* of causation. The first paradigm of causation is a situation in which a person, engaging in ordinary human activities, brings about A in order to bring about B, but he could have prevented B by preventing or not bringing about A. The second paradigm is the converse situation in which a person, engaging in ordinary human activities, prevents B by preventing A, but he could have ensured that B occur by bringing about A or by not preventing A.

While I accept this third account of producing causation, there is a good reason why one requires some explicit account of the features $F_1 \ldots F_m$ on which causation is supervenient, whereas one did not require such an account in the case of tablehood. The reason is that it is not absurd to be a causal realist. Suppose one could not explicitly describe the features $F_1 \ldots F_m$, in non-causal terms (that is, one could not say in what non-causal respects cases of causation are similar to the paradigms), then the causal realist could argue that at least one of the features $F_1 \ldots F_m$ should be described in causal terms. Conversely, if one can explicitly describe the features $F_1 \ldots F_m$ in non-causal terms, then, using Ockham's Razor, one can argue against there being any irreducibly causal element "in the objects". But in the case of tablehood, no one claims that there is an irreducible element of tablehood "in the objects" and so there is no need to explicitly describe the features on

which tablehood is supervenient. Thus neither the Conditional Analysis nor the Regularity Account is a rival anti-realist theory of causation to the Manipulability Account, but rather both supplement the Manipulability Account by describing in what respects causal situations are similar to the paradigms of causation.

The need for both a conditional analysis and a regularity account arises from the possibility of a stricter and a less strict requirement for an account of the features on which causation is supervenient. The stricter requirement is that in *all coherent examples, even hypothetical ones,* A causes B if and only if the features $F_1 \ldots F_m$ are present. The less strict requirement is that, as part of one's metaphysical hypothesis, one might rule out some coherent hypothetical examples as *impossible* (in a non-logical sense of 'impossible'), so one only requires that A causes B if and only if the features $F_1 \ldots F_m$ are present *in actual examples, or in possible hypothetical examples.* Using the stricter requirement, that is, the Analysis-criterion, one obtains the Conditional Analysis which, as I argue in Part Two of Chapter One, provides an account of the features on which causation is supervenient in terms of the history of the world up to the causal situation and in terms of non-accidental generalisations. Using the less strict requirement, that is, the Reduction-criterion, one obtains the Regularity Account which gives an indisputably non-causal description of the features $F_1 \ldots F_m$.

0.4 There are three theses about Metaphysics which my discussion of causation illustrates.

Thesis One: Although ontological theories about what is "in the objects" and what is "in the mind" might be considered rather

vague, the arguments used to discuss such theories can be quite precise and detailed. For, using the method of counter-examples (to be discussed in Chapter One, Section One), one can provide a clear and coherent counter-example to some claims about what is "in the objects". That vague theories are supported by precise arguments is no more surprising than that the longest worm (whose length varies at it moves) is definitely longer than the standard metre.

Thesis Two: An interesting metaphysical theory (such as the Regularity Account of Causation) can be held as part of a metaphysical hypothesis, which, although it cannot be conclusively proved or disproved, can be compared with rival metaphysical hypotheses.

Thesis Three: Even if one is interested in what occurs "in the objects" and one wants to describe what occurs without using anthropocentric terms, one needs to consider *ordinary language*. In the discussion of causation, if there were situations in which, according to the account, A should cause B, but it is not *ordinary usage* to say that A causes B, then one has not succeeded in describing the features $F_1 \ldots F_m$ on which the *ordinary notion of causation* is supervenient, and so one is open to the objection that one of the features $F_1 \ldots F_m$ can only be described in causal terms. Once one has argued for an anti-realist theory of causation one is *then* at liberty to ignore ordinary usage and to stipulate a new use for the word 'cause'.

0.5 My account of causation is based on Mackie's in *The Cement of the Universe*, who provides an excellent discussion both of the Conditional Analysis and of causal regularities.

I assume that causes are no later than their effects until I discuss backwards causation in Chapter Four. In that chapter I argue in favour of two negative theses. On the one hand, there are no good arguments for the logical impossibility of backwards causation. On the other hand, there are no examples which cannot be reasonably interpreted by the causal anti-realist as not involving backwards causation. It follows that no proposed case of backwards causation could be used as a counter-example to an anti-realist account of causation. In particular, the accounts of Chapters Two and Three need no revision in the light of Chapter Four.

CHAPTER ONE

SOME PRELIMINARIES

Section One: The Method of Counter-examples

1.1 As I have indicated in the Introduction to this thesis, my aim is to provide two accounts of transeunt causation: one is based on Mackie's Conditional Analysis of causation; the other is based on his discussion of causal regularities. In this section I expound the method which I propose to use for testing the permissibility of accounts of causation. I assume that in any account of causation, sentences of the form 'X causes Y' are systematically correlated with sentences stated in non-causal terms. For example, one might attempt to give an account of causation by correlating 'X causes Y' with 'X occurs, Y occurs, X and Y are distinct existences, and if X did not occur Y would not occur'. Such a correlation is systematic in the sense that the same correlation is made for a large variety of Xs and Ys, say for any events X and Y. I do not think that anyone would accept an account of causation which is not systematic in this sense. It would, for instance, be arbitrary in the extreme to give a regularity account of causation if cause and effect occurred *before* 1900, but to give some other account if cause and effect occur *after* 1900.

Thus any proposed account of causation should be a systematic correlation of sentences with sentences of the form 'X causes Y'. By 'sentence' I mean a sentence in a natural language rather than marks on paper. For the purposes of this thesis it seems best to consider sentences rather than propositions. I do not want to consider

propositions since it is possible that the sentence 'X causes Y' is not strictly speaking true or false. In that case one can discuss its "truth" or "falsity" using generally accepted pre-philosophical criteria of truth-values, yet the sentence does not correspond to a proposition.

1.2 The Analysis-criterion:

The method which I shall use in examining accounts of causation is the *method of counter-examples*. Suppose, initially, that the purpose of the account is to provide a sentence p - called the *analysans* - which is *synonymous* with a given sentence q - the *analysandum*. Thus p and q are to have the same cognitive meaning. If one can coherently describe a situation in which p would be true and q would be false or vice versa, then p and q cannot have the same meaning; so one has found a counter-example to the account. For example, suppose p is the sentence 'This animal has a heart' and q is the sentence 'This animal has a kidney'. Then I claim one can tell the following coherent story:

> A certain species of insect, with neither heart nor kidney, lives in the adult stage only for twenty-four hours and so in fact requires no organs for eliminating waste. Like all insects it is too small to need a pump to pump its blood. Now suppose that from this species there evolves a species of insect the size of a rabbit which has a pump to pump its blood, but since it still lives only twenty-four hours in the adult stage it requires no kidneys.

In this case, p would be true and q would be false. This example shows how sometimes, by providing sufficient details, one can argue that a proposed counter-example is empirically possible and, *a fortiori*, coherent. Another method of arguing for the coherence of a

counter-example is to argue that the impossibility of the counter-example was discovered *a posteriori* and so is merely a nomic (or some other non-logical) impossibility. For example, an account of material objects in terms of aggregates of molecules could not satisfy the analysis-criterion since it was discovered *a posteriori* that material objects are not made of homogeneous matter. Again, one can sometimes argue that a counter-example is coherent on the grounds that it can be imagined. Thus one can imagine what it would be like if all red objects seemed green and all green objects seemed red; so, it could be argued, hypothetical examples of colour-reversal provide coherent examples (perhaps counter-examples to some analysis of perception in terms of discrimination-behaviour).

If it is assumed that a necessary condition for the correctness of an account, say of 'X causes Y', is that there are no coherent counter-examples in which the two sentences to be correlated differ in truth-value, then I shall say that the analysis-criterion is being used.

1.3 Some Remarks on the Analysis-Criterion

(1) There are accounts which satisfy the analysis-criterion but in which the analysans and the analysandum are not synonymous. In particular, if p and q are any two logical or mathematical truths, an account in which p and q are correlated would satisfy the analysis-criterion. Thus, if Goldbach's conjecture is true, there could be no counter-example in which '2 + 2 = 4' is true but 'Goldbach's conjecture is true' is false. Yet '2 + 2 = 4' and 'Goldbach's conjecture is true' do not have the same meaning.

(2) Furthermore, there might perhaps be sentences p and q such that it is neither a mathematical nor a logical truth that p≡q but such that

an account correlating p and q satisfies the analysis-criterion. (Here, if p and q were synonymous, I would treat p ≡ q as a logical truth.) For example, suppose it is synthetic *a priori* that nothing can be uniformly red and uniformly green at the same time. Let p be the sentence 'This surface is uniformly red' and q be the sentence 'This surface is uniformly red and not uniformly green'. Then p ≡ q is neither a logical nor a mathematical truth; but one could not, I suggest, find a coherent hypothetical example in which p is true and q is false.

To this it might be *objected* that if the sentence 'No surface is uniformly red and uniformly green' is synthetic then it follows that one can coherently describe a hypothetical situation by simply saying that there is a surface which is uniformly red and uniformly green. To this *I reply* that the analysis-criterion, to be useful, requires an independent criterion of coherence. Suppose one argues that p ≡ q is neither logically nor mathematically necessary and so p & ~q is a coherent description of a counter-example to an analysis in which p is the analysandum and q is the analysans. In that case the analysis-criterion is circular and hence of no interest. I stipulate that the analysis-criterion is to be used with examples whose coherence is not asserted *purely* on the grounds that two sentences are not logically equivalent.

(3) The analysis-criterion is not a precise criterion. There are two reasons for this imprecision. The first is that some examples are of debatable coherence. For example, consider the following famous passage from Descartes:

> "I will suppose ... that there is an evil spirit ... I will suppose that ... colours, shapes, sounds and

> and all external objects are mere delusive dreams ... I will consider myself as having no hands, no eyes, no flesh, no blood, no senses, but just as having the false belief that I have all these things." [1]

This example might perhaps be used as a counter-example to an idealist analysis of sentences about material objects. But its coherence is debatable.

The second reason is that in some proposed counter-examples, the example is described coherently but the situation being considered diverges so much from paradigms that there could be dispute over the truth-values of the analysandum and/or the analysans. For example, Dummett considers the following as an example someone might want to give of backwards causation:

> Someone who believes in magic ... has among his spells a formula for producing good weather in a particular place on a particular day; this formula works without fail. ... An occasion arises when he has a reason for wanting the weather at, say, Liverpool, to have been good on the previous day, but he does not know whether there was [good weather] or not; he therefore recites his spell, putting in yesterday's date. Subsequently he finds out that there was fine weather at Liverpool on that day; and he finds that whenever he recites the formula with a past date, not knowing what the weather was like on that date, later investigation proves the weather to have been fine then. [2]

This example is coherent, but it has been disputed whether the magician's spell caused the clear weather the day before.

---

1. Descartes, *First Meditation.*

2. Dummett, M.A.E., "Can an Effect Precede its Cause?" *Aristotelian Society Proceedings,* Supp. Vol.28 (1954), p.35.

However, it would be rash to conclude that the analysis-criterion is useless as a necessary condition for the analysis to be correct. Even if one restricts oneself to using counter-examples, which are clearly coherent, and in which analysans and analysandum have truth-values assignable without controversy, one can find counter-examples to many initially plausible analyses. For example, I claim[3] that one can coherently describe a world in which one would say of two events A and B that A caused B but no event like A ever occurred on any other occasion. This counter-example prevents one treating certain regularity accounts of causation as analyses.

(4) The analysis-criterion is nevetheless interesting even if one rejects, like Quine, the analytic-synthetic distinction. Suppose one accepted Quine's conclusion:

> Any statement can be held true come what may, if we make drastic enough adjustments elsewhere in the system ... Conversely, by the same token, no statement is immune to revision. Revision even of the logical law of the excluded middle has been proposed as a means of simplifying quantum mechanics. [4]

It does not follow that people are able to treat at will any truth as synthetic. The available range of conceptual schemes or systems is determined by culture, language, and available scientific theories, or even by human nature. Hence the task of discovering what pairs of sentences are treated as equivalent would still be an important part of the study of conceptual schemes. As conceptual schemes change so, presumably, would the judgments about the truth of sentences in coherently described situations.

---

3. See pp.165, 166 for the details of this example.

4. Quine, W.V.O., "Two Dogmas of Empiricism", in *From a Logical Point of View*, p.43.

(5) Presumably there are other necessary conditions for an analysis, in addition to satisfying the analysis-criterion and in addition to the requirement that the analysis be systematic. For example, if p is the analysandum and q is the analysans, there is another candidate for an analysis in which p is the analysandum and p & q is the analysans. This would be systematic and it would satisfy the analysis-criterion. But surely this is not a permissible analysis. In some way, the analysis has to result in some clarification or simplification. However, I shall ignore such considerations which are hard to state formally; when I am discussing the analysis of causation I shall consider only the analysis-criterion.

1.4 The Reduction-Criterion

An account which fails to satisfy the analysis-criterion might still be important. For example, a regularity account of causation - which fails to satisfy the analysis-criterion - might well be the correct account of causation "in the objects". Again, suppose one makes the *contingent* identification of material objects with aggregates of molecules, and one *contingently* identifies the colour of the surface of a material object with the property of emitting light of certain wave-lengths. In that case, one obtains an account of the sentence 'This material object is red' as 'This aggregate of molecules emits light of wavelength of 7,000 to 8,000 Angstrom units'. This account is not an analysis, yet the scientific realist claims the account gives, in some sense, a better or more fundamental description of some state-of-affairs. This is admittedly somewhat vague. Yet I claim the vague notion of providing an account of what occurs "in the objects" or of providing a more fundamental account, is clear enough to provide one

with a new criterion in place of the analysis-criterion. Therefore in my discussion of a regularity account of causation I shall use this new criterion without either considering what it means to provide an account of what occurs "in the objects", or considering what it is to provide a more fundamental account.

Such accounts I shall call, for want of a better term, *reductions*. What criterion should one have for reduction? A necessary condition for any account of what occurs "in the objects" is surely that there be no *actual example* in which reductiandum and reductians have different truth-values. For example, suppose the reductiandum is 'X causes Y' and the reductians is a regularity account. The claim is that when someone says 'X causes Y', and according to the usual criteria he is correct, whatever he thinks might be happening, all that is actually happening "in the objects" is that there is an instance of some regularity. If on some actual (not hypothetical) occasion, 'X causes Y' is true but the reductians is false, then one has incorrectly described what is occurring "in the objects". Conversely, if the reductians is true but X does not cause Y, the account is incomplete; one has perhaps discovered *some* of what occurs "in the objects" in cases of causation, but, presumably, one has omitted *part* of the description of what occurs "in the objects".

Now this criterion that truth-values are preserved could be used in two different ways. Suppose there is a *proponent* of a reduction and an *objector* to that reduction. *On the one hand* if one assumed that the *objector* was required to find actual counter-examples to the proposed reduction, then the *proponent* of the reduction would benefit from the objector's human limitations. So the proponent would be able to put forward excessively strong reductions. For

example, if the universe is large enough, there are probably regions of the universe in which *accidental* regularities occur which an observer would rationally believe to be causal or non-accidental regularities. Yet an objector to a regularity reduction of causation obviously could not, when challenged, provide an example of a regularity which is accidental and yet which is rationally believed to be causal. *So I reject this use of the criterion.*

On the other hand, if the proponent of the reduction has to demonstrate that there are no actual counter-examples to the reduction anywhere, he would have to argue that such counter-examples are impossible - perhaps in some non-logical sense of 'impossible'. However, a metaphysician is likely to argue for the impossibility of a kind of example on the grounds that it is inconsistent with his metaphysical theories. For example, if someone claims that nowhere is there a material object made of homogeneous matter, his claim is part of a metaphysical theory and is not, say, based on straightforward induction from known to unknown material objects. For an operationalist could believe that in fact material objects are composed of homogeneous matter but that the properties of homogeneous matter are conveniently described in terms of fictions called molecules. Thus it is plausible that the proponent of a reduction which is not an analysis can only defend his reduction as part of a general metaphysical hypothesis. I shall adopt the following procedure, which I shall call the *reduction-criterion:*

> A reduction is *demonstrably incorrect* if an *actual* counter-example can be found. In addition, the proponent of the reduction is required to incorporate into his metaphysical hypothesis some principle or some assertions from

which it follows that no counter-example occurs anywhere at any time. In the event that ad hoc or totally implausible assertions are incorporated into the metaphysical hypothesis one can *reject* the reduction. If the assertions incorporated into the metaphysical hypothesis are fairly plausible, then one has no objection to the reduction, but one has merely revealed the *commitments* of a metaphysician who proposes the reduction.

Note: In *Truth, Probability and Paradox* Mackie considers what he calls factual analysis. He says "There is then, a problem about the factual analysis of causation; we have the task of finding out what goes on in the world in those sequences and processes that we mark off as causal."[5] The factual analysis is an answer to the question "Is there anything common to and distinctive of the members of this class?"[6] One might interpret Mackie's notion of factual analysis in such a way that if by accident there are no counter-examples to the factual analysis then it is correct. However, suppose that it were the case that (by accident) men and only men were, to use Russell's example, featherless bipeds. Would a factual analysis of a man as a featherless biped be acceptable to Mackie? I am not sure. But if some kind of non-arbitrariness or non-accidental character is required for an account to be a factual analysis, then, I suggest, Mackie's notion of factual analysis is rather close to what I call reduction.

---

5. Mackie, J.L., *Truth, Probability and Paradox*, p.12.

6. *Ibid.*, p.12.

1.5 Two Examples of the use of the Reduction-criterion

Example One: Action at a Distance

Suppose that an account of causation is proposed, and if there is action at a distance the reductiandum is true but the reductians is false. Suppose also that the proponent of the account accepts that action at a distance is coherent. Then the proponent of the reduction is committed to include in his metaphysical hypothesis some claim which entails that there is no action at a distance. Now it would not be surprising if there was no action at a distance. Nonetheless, the incorporation of the no action at a distance principle into a metaphysical hypothesis is, I suggest, a defect - though not a great one - in the theory. Other things being equal, one would prefer a metaphysical hypothesis that did not contain as an unargued assumption the principle that there is no action at a distance. For one would like metaphysical hypothesies to be compatible with all kinds of scientific theory which one might expect scientists to produce in the future. Thus while a reduction of causation which commits one to the denial of action at a distance would not be absurd or demonstrably incorrect, it would be less attractive than a simple modification of that reduction without this commitment.

Example Two: Causation and Recipes

On Gasking's account of causation "A statement about the cause of something is very closely connected with a recipe for producing it, or for preventing it."[7] Suppose (as Gasking does not) that one is always to give a straightforward recipe account of transeunt causation. That is, consider the reduction in which the reductiandum is 'X causes Y' and the reductians is 'Events of the Y sort can be produced by

---

7. Gasking, D., "Causation and Recipes", *Mind*, Vol.64 (October 1955), p. 483.

producing events of the X sort'. One might object to this account on the grounds that it is conceivable that the Solar System was caused by the approach of another star near the Sun.[8] More precisely, suppose that another star approaches the Sun and that later (as in a now discredited hypothesis about the origins of the Solar System) a cigar-shaped cloud of gas and dust is dragged away which condenses into the planets. Assume that this example is coherent and also assume that the *proponent* of the reduction *admits* that in this situation one would say that the approach of the star *caused* the Solar System to come into existence. In that case the proponent of the reduction would be committed to the non-occurrence of such events as part of his metaphysical hypothesis. If he could not derive the non-occurrence of such events from more plausible premises, his hypothesis would, I suggest, contain an *ad hoc* assertion; so one would reject the reduction. Perhaps the proponent of the reduction would claim that God produced the Solar System by making a star approach the Sun. In that case the *ad hoc* character of the metaphysical hypothesis is lessened, but the metaphysician has now committed himself to an account in which God produces one event in order that another event should come about; the objector to the reduction could then set about examining the peculiar consequences of the proponent's commitment.

---

8. This example is Mackie's. See Mackie, J.L., *The Cement of the Universe*, p.170.

Section Two: Events and Causal Fields

In Section Two I consider the notions of a *quasi-event* and of a *causal field*. The notion of a quasi-event is required to make precise the category to which cause and effect belong in cases of producing causation. The notion of a causal field is required for two purposes. First, it enables one to define quasi-events. Second, the claim that causation is a *three-term* relation between *cause*, *effect* and *causal field* enables one to give an account of the peculiarity of some causal sentences. I want to treat these peculiar sentences as false in order to keep the notion of causation which I am analysing fairly close to those *producing paradigms*, in which a person produces the cause as a means to producing the effect - and not merely to the "*preventing paradigms*" in which one could have prevented the cause by preventing the effect. The features which in ordinary circumstances no one ever has to produce because they are *always* present are thus not treated as causes. For example, there is something peculiar, I claim, about the sentence 'The presence of oxygen in the atmosphere was a cause of the fire'. There is a broader notion of causation, departing further from the paradigms, in which one *would* say that the presence of oxygen was a cause of the fire. This broader notion can be obtained by considering the *null causal field*; in that case *almost* any fact about what occurs in some region of space-time can be treated as a quasi-event.

I also mention causal fields when I discuss the notion of distinct existences and the notion of a regularity.

2.1 My thesis is concerned only with transeunt causation - not with immanent causation. The paradigms of transuent causation

I take to be cases in which one event causes another event. However, if one asked 'What caused the fire?' one might accept as equally appropriate the answer 'Dropping a cigarette', which describes an *event*, and the answer 'A fault in the wiring' which does not describe an event but rather a *condition*. This suggests that cases of causation in which the cause is an event might form an artificially restricted class. So I shall not rule out cases of transeunt casuation in which the cause is not, strictly speaking, an event. However, someone might argue that one can always analyse sentences about causes in terms of sentences about events. Thus one might attempt to analyse 'The faulty wiring caused the fire' as 'The occurrence of a fault in the wiring caused the fire'. Davidson suggests that examples of causal idioms which resist such an analysis are rudimentary *causal explanations*[9]; he gives the example "The collapse was caused not by the fact that the bolt gave way, but by the fact that it gave way so suddenly and unexpectedly". It seems that Davidson would claim that the cause of the collapse was the event of the bolt giving way, and the purpose of this rudimentary causal explanation is to point to some feature of the covering causal law. In this case, even if we cannot fully state the covering law it would be of the form: 'If e is an event ... which is rapid enough and occurs at time t then e is followed at time t+ε by an event ....' Thus the properties of events are considered in the description of the causal law rather than in the description of the cause itself. However, the difficulty with Davidson's account is that there are cases of transeunt causation in which there is no event which is

---

9. Davidson, D., "Causal Relations", *The Journal of Philosophy*, vol.64, (Nov. 1967), pp. 702-703.

identifiable as the cause. For example, consider the question 'What caused the fire to burn out of control?' An acceptable answer might be 'The use of a water-based fire extinguisher'. An equally acceptable answer would be 'The presence of pure oxygen'. In this latter case what is the event which caused the fire to *burn out of control?* Surely not the same event, say dropping a cigarette, as caused the fire to *burn*, since 'Dropping a cigarette' would not be the correct answer to 'What caused the fire to burn out of control?' Nor need it be the event of the oxygen coming to be present, for there might *always* be pure oxygen in that room. On Davidson's account one seems to have a causal explanation without any cause.

Hence there seem to be cases of transeunt causation in which a fact rather than an event is a cause. Indeed there is an idiom in which 'A causes B' might be analysed as 'A and B are logically independent and A explains B'.

I shall now argue that the notion of transeunt causation is based on two different kinds of paradigm of causation; producing paradigms and explanatory paradigms.

In many cases in which one can *produce* Y by producing X, the occurrence of X *explains* the occurrence of Y. *For example,* if a short-circuit causes the fuse to blow, one could have made the fuse blow by producing the short-circuit. One can also explain the blowing of the fuse in terms of the short-circuit.

However there are some cases in which a producing paradigm is not an explanatory paradigm. *For example,* suppose that it is discovered empirically that an extract of jojoba beans sends people to sleep. Suppose that it is not known what chemical compound in the

extract is responsible for this property, *and* that it is not known how the extract acts on the central nervous system. This, I claim, is a *producing paradigm*; people are sent to sleep by being given extracts of jojoba beans. No doubt, in this case, one might *explain* the person going to sleep by saying that he had taken an extract of jojoba beans. But, I claim, the explanation depends on the *prior* knowledge that jojoba bean extracts can be used to produce sleep. So this is not an explanatory *paradigm*.

Conversely, there can be explanations which do not closely resemble producing paradigms. Suppose, for example, that a "mechanical tortoise" moves around and that one is perplexed by its movement until one notices that, when it reaches an electric power-point, it recharges its batteries. In that case, one has *partially* explained the motion of the mechanical tortoise in terms of the event of its recharging itself. Yet one cannot prevent or produce that motion by preventing or making the tortoise recharge itself. One is reluctant even to say that the tortoise recharging itself *causes* its (earlier) motion. Rather, one tends to assume that there is some mechanism in the tortoise which causes its power-point-seeking behaviour. But the tendency to assume that there is such a mechanism is based on the prior success of the explanation of the motion of the tortoise in terms of its recharging itself. If, on the one hand, one *does* say that the tortoise recharging itself *causes* its (earlier) motion then one has an example in which there is an explanatory paradigm which is not a producing paradigm. If, on the other hand, one does *not* say that the tortoise recharging itself *causes* its (earlier) motion then it seems that the notion of transeunt causation is based more on producing paradigms than on explanatory paradigms. Instead of settling this

dispute over the use of the term 'cause', I shall distinguish two kinds of transeunt causation; explanatory and producing causation. Explanatory causation is based on the explanatory paradigms; producing causation is based on the producing paradigms. I shall concentrate, in this thesis, on producing causation. Thus I shall investigate the features in virtue of which cases of transeunt causation resemble producing paradigms, and I shall largely ignore any features in virtue of which some cases of transeunt causation - it might be claimed - resemble those explanatory paradigms which are not also producing paradigms. However, in Chapter Four (Section 3.5) I shall discuss why (producing) causes *seem* to explain their effects not vice versa, in order to answer an objection to Causal Anti-realism.

2.2 Causal Fields

In *The Problem of Causality*[10] Anderson claims that the cause is a differentia of some genus and the effect a variable property of members of the same genus. The genus is the causal field. Anderson says

> The inquiry into causes, ..., is only a special case ... of the solution of problems in general. In trying to determine when a phenomenon is present, and when it is absent, in a given field, we are endeavouring to divide a genus (the field) into two species, one of which has a certain property, while the other has the opposite.[11]

Thus to take Anderson's example, if one asks for the cause of my anger, the anger of men or the anger of animals in general, one is considering three different causal fields and so one should expect different answers.

---

10. Anderson, J., "The Problem of Causality", *Australasian Journal of Psychology and Philosophy*, vol.16 (Aug.1938), pp. 127-142.

11. *Ibid.*, p.134.

Anderson's theory of the causal field does not enable one to discover to what category causes and effects belong *until* one has discovered to what category the causal field belongs. In many cases both cause and effect are changes in some objects, so the causal field might be the existence of and persistent properties of these objects. However, if the effect is a fire, it is the *destruction* of an object in a certain manner rather than a change. Such examples suggest that a causal field is a genus whose species consist of certain states of affairs in various regions of space-time. More generally, to cover the case of determinables, it seems that Anderson's causal field consists of a range of possible states of affairs in regions of space-time. For example, one says 'The spark caused the fire' but, I claim, there is something peculiar about 'The presence of oxygen in the atmosphere caused the fire'. This is because all the states of affairs in the range being considered are ones in which oxygen is present, but not all are ones in which there are sparks.

The difficulty with Anderson's account of the causal field is in determining, in any given case, what the causal field is. Anderson himself seems to have considered that one should specify a causal field when asking what caused a certain event; he claims that "only confusion can result, if we have not begun by specifying (a) the field, (b) the phenomena ...".[12] In the manner of Anderson I shall adopt the principle that the specification of a causal field is either stated or presupposed when one asks 'What caused B?' or 'What will A cause?' It follows that there is a weighty objection to this account of causal fields; a completely unexpected answer might be quite acceptable. For example, if I ask 'What caused the

---

12. Anderson, J., *op.cit.*, p.134.

house to burn?' I have in mind a range of possibilities. In particular, I assume that the atmosphere consists of ordinary air, but the answer 'The atmosphere in Hobart was, for two hours, 100% oxygen' would, if true, be an acceptable answer.

There might be several ways of avoiding this difficulty. I choose the following. The causal field is not a range of possibilities, but rather a norm or standard with which the actual situation is compared. Thus, when one asks 'What caused the house to burn?', one presupposes the normality of an atmosphere containing some oxygen, but not the normality of an atmosphere of pure oxygen. Hence the presence of oxygen in the atmosphere is not a departure from the norm, but the presence of 100% oxygen in the atmosphere is a departure from the norm. So the latter - but not the former - is a candidate for the cause of the fire. I do not assume that there are objective criteria for when a state-of-affairs is normal. I merely assume that people make judgments about what is normal. These judgments could vary from context to context. Strictly speaking, there is a whole *class* of normal states-of-affairs - the norms; a cause or an effect can be a departure from *any* member of this class. For example, if the speaker assumes that Joe always drives home by either route five or route ten then there could be two normal states-of-affairs; in one, Joe drives home by route five, in the other by route ten. In that case 'Joe's driving home by route five' might be an acceptable answer to 'What caused Joe to be late?', because Joe's driving home by route five differs from one of the normal states-of-affairs.

Accordingly, I define *the causal field* to be the class of all states-of-affairs which are not considered to be abnormal; I define *quasi-event* to be a state-of-affairs which differs from *at least one*

of the normal states-of-affairs in the causal field. Three points should be noted about these judgments of normality:

(1) Since there can be several normal states-of-affairs, a state-of-affairs can be judged moderately unusual and still be normal in the required sense.

(2) In some contexts, in particular in scientific contexts, very little is deemed *abnormal*. (However the state-of-affairs in which in the year 2000 A.D. all blue objects turn green and vice versa would still be considered abnormal.) The resulting causal field I call the *null field*.

(3) Judgments of normality are made on background evidence only; one does not take into consideration knowledge of the effect or the cause when judging what is normal.

If A is a quasi-event, A is not usually the total state-of-affairs. A is merely some fact or facts about what is occurring in some region R of space-time. Contained in and perhaps identical to R will be a region D of space-time in which A differs from some normal state-of-affairs. D is the *spatio-temporal location* of A.

If A and B are two quasi-events, A is said to be strictly earlier than B and B is said to be strictly later than A if the locations of A and B do not overlap and any point in A is earlier than any point in B.

Someone might argue that the class of quasi-events should be restricted so that excessively remote regions of space-time, or excessively minute details, are not considered. Thus an elaborate account of the evolution of a species of virus and of the evolution of *Homo sapiens* would not, for instance, be acceptable as part of an answer to

'What caused the 'flu epidemic?' However, in this case there are other, shorter, answers which are more appropriate in most contexts; it is a matter of *conversational point* that one does not give excessively detailed answers where simpler ones satisfy the curiosity of the questioner.

The criterion for discovering the causal field presupposed by the question 'What caused B?' or the question 'What will A cause?' is to decide what answers would be considered unacceptable or peculiar, and why. Some answers are peculiar because of considerations of *conversational point*. Other answers are unacceptable because they are part of the description of every normal state-of-affairs in the causal field.

2.3 Two objections to the Notion of a Causal Field

In this subsection I reply to two objections to the use of the notion of a causal field in discussing causation.

Objection One: There is no need to use the notion of a causal field when analysing causation. Rather, one should distinguish between the *conversational point* of a sentence and the *truth-conditions* of that sentence. Consider the example in which it is said to be false that the presence of oxygen caused the house to burn. In this example it is equally peculiar to say 'The presence of oxygen was a cause of the house burning' and to say 'The presence of oxygen was *not* a cause of the house burning'. It is not that it is *false* that the presence of oxygen is a cause, but rather that there is *no point* in saying that the presence of oxygen is a cause. Typically one does not

make assertions that are uninformative or uninteresting. A condition that is always present is not of much interest as a cause, so one rarely, if ever, says 'The presence of oxygen caused the fire'. From lack of use the sentence seems peculiar.

Reply: The following five considerations should meet this objection. However, I admit that in many cases considerations of conversational point could replace or supplement the reference to causal fields.

(1) I claim that both the sentence 'The fire was caused by the presence of oxygen in the atmosphere' and its internal negation make the false presupposition that the phrase 'the presence of oxygen in the atmosphere' refers to a particular in the correct category for a producing cause. For the phrase 'the presence of oxygen in the atmosphere' does not describe a quasi-event (relative to a causal field in which oxygen is present in the atmosphere in all normal states-of-affairs).

(2) What rules are there for deciding when it would be *pointless* to utter a sentence? Considerations of what the hearer *knows* do not explain why some causal sentences are peculiar. For, presumably, there is point in reminding someone of something he knows but of which he is not explicitly thinking. For example, Jane might remind Harry that it is one o'clock and that his train leaves at two o'clock, even though Harry knows perfectly well what the time is and when his train leaves. In the same way, there might be point in reminding someone that oxygen is a cause of the fire, provided he is not thinking of the fact that oxygen is a cause of the fire. Thus I suggest that some criterion of the

relevance, importance, or interest of the sentence would have to be used in order to show that it is, usually, pointless to say 'The presence of oxygen caused the fire'.

(3) There are cases in which 'X was a cause of Y' would be a pointless utterance, yet 'X was a cause of Y' does not seem peculiar in the way that 'The presence of oxygen in the atmosphere was a cause of the fire' seems peculiar. For example, consider the sentence 'Striking the match caused it to light'. Everyone knows that one strikes matches to make them light. Of course in this example the phrase 'striking the match' refers to an event, not a condition. But why should that affect considerations of conversational point?

(4) The notion of the causal field enables one to distinguish between the non-occurrence of a quasi-event and the occurrence of a negatively described quasi-event. Thus suppose someone is poisoned and fails to take the antidote - a weed growing in his garden - which in fact he does not know is an antidote and which is never eaten. In that case one would not say that the failure to take the antidote was cause of the man's death. For the failure to take the antidote would be part of every normal state-of-affairs, and so is not a quasi-event but is merely the absence of a quasi-event. But now suppose that a man works in an area where it is common to be bitten by deadly snakes. Suppose also that all people working in this area are trained in the use of the antidote and provided with the antidote. In that case, if a man is bitten, if he does not use the antidote and if he dies, one would say that his failure to use the antidote was a cause of his death. For the failure to use the antidote is

a difference between the actual and at least some of the normal states-of-affairs.

Presumably, if one replaces the notion of the causal field by considerations of conversational point, one claims that in *all cases* the failure to take an antidote is a *cause* of death, but only in *some cases* is there any *conversational point* in mentioning that cause of death. One might, perhaps, claim that one is influenced by the procedures of law-courts in finding legal responsibility; consequently, there is point in mentioning a failure to take the antidote only when someone would reasonably be *blamed* for the failure to take the antidote.

Now consider the following example. A chemist discovers a cheap non-toxic chemical compound called *Diazoquinoline*. Diazoquinoline when sprayed on plants, enables them to withstand a great degree of drought. This compound has never been used commercially. The chemist passes a field of wheat - obviously suffering from drought - and is asked 'What caused that crop to fail?' He replies (correctly) 'The lack of rain caused the crop to fail, but if the farmer had used Diazoquinoline he would have saved his crop'. I suggest that the chemist would *not* reply 'The lack of rain and the failure to use Diazoquinoline *caused* the crop to fail', unless he believed that it was the practice of farmers to use Diazoquinoline. In this case, the remark 'If the farmer had used Diazoquinoline he would have saved his crop' is informative, interesting, and relevant, and so also, I suggest, would be the remark 'The failure to use Diazoquinoline *caused* the crop to fail'. Hence, if it is a matter of conversational point that normal occurrences (here the failure

to use Diazoquinoline) are not mentioned as causes, there seems to be no explanation for this principle in terms of the requirement that an assertion be informative, interesting and relevant. Furthermore, if the absence of an event is always considered a cause when that absence could sometimes be considered a cause, then it seems that in almost any causal situation the absence of intervention by agents (human beings or, perhaps, God) would have to be considered a cause. Yet this proliferation of negative causes seems counter-intuitive.

(5) I propose, as a criterion for distinguishing conversational point from truth-conditions, that one considers answers to questions in a questionnaire. Assuming that one is co-operative, one answers *yes/no* questions in a questionnaire even if the question seems pointless. I suggest that the perplexity felt on being asked 'Was the presence of oxygen (in normal concentrations) a cause of the fire?' is a sign that, with respect to commonly accepted causal fields, the question is based on a false presupposition, namely that the presence of oxygen is the kind of particular which could be a producing cause. If it is merely a matter of conversational point, one would write 'Yes' with irritation, perhaps, but not with perplexity.

Objection Two: To make causation relative to a causal field is to commit oneself to the claim that causation is to some extent *subjective*.

Reply: To the extent that the causal field depends on what the speaker assumes to be normal, causation is subjective. The following three considerations should make this conclusion more palatable.

(1) Suppose two people are arguing at cross-purposes because they have in mind different causal fields. If they are aware of the way in which causation is relative to a causal field they are then in a position to reach agreement on the choice of a causal field. Moreover, there seems to be a convention whereby if someone says 'X causes Y' the hearer assumes a causal field is being considered such that the utterer of 'X causes Y' does not make false presuppositions. For example, suppose that in answer to the question 'Why did the grass grow so well this winter?' Jim says 'The rain caused the growth'. Suppose that Anne had thought that the unusually warm weather caused the growth and Anne *assumes* that winters are wet. Then Anne would infer that Jim had in mind a different causal field from hers, namely one in which reliable winter rain is not considered part of every normal state-of-affairs. Hence I conclude that the subjective nature of causation in no way lessens the usefulness of sentences about causation.

(2) In order to obtain an *objective* notion of causation one can stipulate that one is considering the *null* casual field. The sentence 'Relative to the null field, the presence of oxygen was a cause of the fire' *is* true. There is, however, a good reason why one does not *always* consider the null field. For, relative to the null field, there are *too many* causes. Here one could, of course, appeal to considerations of conversational point to limit the range of mentioned causes. But I have argued that this is not, in fact, the practice of speakers of English.

(3) It is important to distinguish between a subjective *relatum* and a subjective *relational sentence*. For example, someone might claim that the ineffable content of coloured sense-data is subjective in the sense that it varies from person to person. Nonetheless, the

relational sentence 'The sense-datum intuited when someone looks at a ripe tomato differs in hue from the sense-datum intuited when someone looks at a lettuce' would, presumably, be objective.

If, as I claim, there is a three-term relation between causal field, cause, and effect, the *relational sentences* are *objective* in the following sense: given a causal field, a proposed cause, and a proposed effect, if two people differ in their judgments about whether the relation holds, at least one person is mistaken. However, one of the *relata* - the causal field - is (in many cases) *subjective*, in that it depends on what the person *judges* to be *normal*. Thus the *two-term relational sentence* 'X is a cause of Y' as well as one of the *relata* for the three-term relation are *subjective*, but the *three-term relational sentence* 'Relative to F, X is a cause of Y' is *objective*.

2.4 The Logic of Quasi-events

A quasi-event is a state-of-affairs differing from at least one norm. Two quasi-events are said to be *equivalent* if the only difference between them is part of every normal state-of-affairs. Thus 'Joe's taking route five' and 'Joe's not taking route ten' are descriptions of *equivalent* quasi-events if it is part of every norm that Joe takes either route five or route ten.

One can apply truth-functional connectives to quasi-events only if the resulting state-of-affairs differs from some norm. For example, if it is part of every norm that one but not both of events C and E occur, then the occurrence of C and the occurrence of E are individually quasi-events but they have no disjunction.

I shall say that quasi-event A *contains* quasi-event B, if A

is *equivalent* to B & C for some quasi-event C.

Next I shall consider the problem of when two quasi-events are distinct existences. It is assumed that if A causes B, A and B are distinct existences. If A causes B, A and B must not be equivalent. Moreover, it seems that A should not contain B. For example, one would not say 'The burning of the house caused the burning of the bedroom'. Again, the effect does not contain the cause. One would not say 'The burning of the bedroom caused the burning of the house', but rather 'The burning of the bedroom caused the burning of the remainder of the house'. Thus a necessary condition for A and B to be distinct existences is that A does not contain B and B does not contain A.

Furthermore, it seems that cause and effect do not even *overlap*. Thus, while one might say 'Last year's disasters caused this year's disasters', one would not, I think, say 'Last term's disasters caused this month's disasters' where some of the disasters occurred both in last term and this month.

However, it is not obvious what is meant by the overlapping of quasi-events. For, if A and B are quasi-events, A is the same as $(A \vee B) \,\&\, (A \vee \sim B)$ and B is the same as $(A \vee B) \,\&\, (\sim A \vee B)$, so A and B have an overlap, namely $A \vee B$. One might be tempted to consider spatio-temporal overlaps, but there seems to be no prohibition on cause and effect occupying the same region of space-time. For example, one could say that salt being at a temperature of 1000° causes it to emit yellow light. Here both cause and effect occupy the same region of space-time.

I propose a somewhat more complicated account of over-lapping,

based on a claim that (relative to the causal field) one can recognise *disjunctive quasi-events*. If N is any normal state-of-affairs in the causal field and A a quasi-event, there might be a possible state-of-affairs, unique up to equivalence, obtained by modifying N so that A occurs, but no quasi-event occurs which is neither contained in N nor contained in A.[1] In that case, one can call this modified state-of-affairs the product of N and A, written 'N × A'. One can recognise disjunctive quasi-events, since if Y is disjunctive there will be some normal state-of-affairs N such that there is no product N × Y. For if one modifies N so that A ∨ B occurs, and both A and B are departures from N, then there is in general no state-of-affairs obtained from N such that A ∨ B occurs but not such that either A occurs or B occurs. Here I make a proviso about what is *not* specified in the description of hypothetical states-of-affairs. One can consider a hypothetical state-of-affairs in which certain kinds of detail are not specified and so the corresponding sentences have *truth-gaps*. But such a lack of detail must be systematic. For example, if it is not specified whether one frog is made up of cells, then it should not be specified whether any frog is made up of cells. This consideration, though somewhat vague, enables one to reply to the objection that N × (A ∨ B) might be such that the sentences 'A occurs' and 'B occurs' have no truth-values. Of course, in some cases this would happen – for example, if A is equivalent to C & D and B is equivalent to C & ~D where D and its negation ~D involve excessive detail.

I now define *regular* parts of quasi-events as follows. A quasi-event B is a regular part of quasi-event A if there is some quasi-event C such that, for every normal state-of-affairs N:

(1) N × A is defined;

(2) $(N \times B) \times C$ and/or $(N \times C) \times B$ are defined;

(3) $N \times A$ is equivalent to $(N \times B) \times C$ and to $(N \times C) \times B$, or if only one is defined to whichever is defined.

Accordingly, I claim that the following conditions are individually necessary and jointly sufficient for A and B to be distinct existences.

(1) A does not contain B;

(2) B does not contain A;

(3) A and B have no *regular* part in common.

I now consider an example which might *seem* to be a case in which the above account of distinct existences fails.

Example:[13] Joanne has a child and so Harry becomes a father. Yet one does not say that Joanne's having a child *caused* Harry to become a father.

One is tempted to say that it is part of the causal field that if Joanne has a child it is Harry's. But, of course, it might be considered a departure from a normal state-of-affairs that Harry rather than Tom or Dick was the father. This example, I suggest, incorporates some of the features of *collateral effect* situations. Here is suffices to say that Harry becoming a father is a quasi-event contained in the conjunction of Joanne giving birth and all earlier quasi-events.[14]

---

13. Based on an example due to Kim. See Kim, J., "Causes and Counterfactuals", *Journal of Philosophy*, vol.70, (Oct.1973), p.571.

14. See pp.87-8 for the details of this modified distinct existences requirement.

2.5 Causal Fields and Regularities

A regularity is a universal (or almost universal) generalisation with, I assume, at least one instance. Usually a regularity has many instances. The typical form of the regularities which I consider in Chapter Three is:

> Whenever a quasi-event *qualitatively identical* to A occurs then a quasi-event *qualitatively identical* to B occurs.

The notion of a regularity is related to that of a causal field, because one can define qualitative identity *relative to causal fields* as follows:

> Quasi-events A and B are qualitatively identical *relative to the causal field F* if they have descriptions differing only in time and place and without reference to other quasi-events such that *for all times and places* that description is of a possible quasi-event.

Example One: If the description is 'A house burns at position x and time t', then for no position x and time t is it part of every norm that a house burns at position x and time t. Hence the quasi-event A that a house burns at position x_1 and time t_1 is *qualitatively identical* to the quasi-event B that a house burns at positive x_2 and time t_2.

Example Two: The description 'An emerald turned grue at position x and time t' is not allowed when discussing qualitative identity, since with respect to all usual causal fields, it does not describe a quasi-event if

t is the year t_o in which everything bleen turns grue. (The predicates 'grue' and 'bleen' are those of Goodman's New Riddle of Induction.[15])

The second example illustrates the convenience of considering qualitative identity *relative to a causal field.* By considering the causal field one avoids treating the event that an emerald turns blue (in some year t_1 before t_o) as qualitatively identical to the "event" that an emerald stays green (in the year t_o). The New Riddle of Induction can be seen as a problem about the way causal fields are chosen, that is a problem about the judgments of normality which people make. The New Riddle of Induction does not threaten the Regularity Account of causation, since both causation and regularities are relative to the same causal field.

Note One: One might, in a regularity, specify the spatio-temporal relation between the correlated quasi-events. I shall ignore the spatial relation but, typically, I shall consider regularities in which quasi-events identical to A are *followed* by quasi-events identical to B.

Note Two: Since there can be disjuncts and conjuncts of quasi-events, a complex regularity can be put in the form given above.

Note Three: In Chapter Three, I shall consider *time-dependent* regularities.[16]

---

15. See Goodman, N., *Fact, Fiction and Forecast,* pp. 73, 74.

16. See pp. 129-30.

Section Three: The Conditionals Used in the Analysis of Causation.

In Part One of Chapter Two I shall discuss a conditional analysis of causation, which involves counterfactuals such as 'If A had not occurred, B would not have occurred'. In Part Two of Chapter Two I shall give an account of the conditionals used in that analysis. I avoid giving a general account of counterfactuals, rather I consider five accounts of counterfactuals and I use these accounts to provide *stipulative* definitions of kinds of counterfactuals.

I assume that four of these kinds, namely *condensed argument conditionals, meta-inference conditionals, supposition/assertion conditionals* and *rationality-judgment conditionals* are genuine cases of conditionals and so are candidates for the conditionals used in the analysis of causation. But I do not assume that all conditionals are of these four kinds.

I avoid the question of whether Lewis-conditionals are *genuine* conditionals and I only consider them in the context of Lewis' own account of causation. I do *not* consider Lewis-conditionals to be candidates for the conditionals used in my conditional analysis. This reflects my bias against an *ontology of possible worlds*. I require the features on which causation is supervenient to be features of the *actual world*.

Note: I am *not* attempting to provide a detailed discussion of various accounts of conditionals. My purpose in discussing these five kinds of conditional here is merely to simplify my discussion, in Part Two of Chapter Two, of the kind of conditionals to be used in the Conditional Analysis. Any details of the accounts of conditionals that affect the discussion in Chapter Two will be discussed in

that chapter.

(1) Condensed Argument Conditionals and Meta-inference Conditionals

In "Counterfactuals and Causal Laws", Mackie puts forward a condensed argument account of counterfactual conditionals. Counterfactuals were, he claimed, condensed and incomplete arguments. Thus he said of the counterfactual 'If you had had the brakes fixed there would not have been a collision' that it might be expanded to 'Suppose that you have had the brakes fixed. Then when the other car turns across your path you press the brake pedal. So your car stops quickly. So there is no collision'.[17]

In any analysis of *ordinary* counterfactuals, it is important to stress the *incompleteness* of the argument, that is, the need for extra premises. However, if one is analysing causation in terms of counterfactuals it seems best to explicitly state the missing premises. Therefore I stipulate that 'If p then q' is a condensed argument conditional, if it can be expanded as a condensed *complete* argument: 'Suppose p, then ... q'. In the analysis of 'Quasi-event A caused quasi-event B' I shall use the conditional 'If, under the circumstances, A had not occurred B would not have occurred'. The phrase 'under the circumstances' reminds one of the extra premises required in the analysis.

In a discussion of the meaning of counterfactuals, it is necessary to distinguish the condensed argument 'Suppose p therefore q' from its metalinguistic counterpart ' 'q' can be inferred from 'p' '. The former is a better candidate for the *meaning* of counterfactuals

17. Mackie, J.L., "Counterfactuals and Causal Laws" in *Analytical Philosophy*, ed. Butler, R.J., p.68.

since one would say that counterfactuals are not about sentences or inferences. However, for the purposes of the analysis of causation, the metalinguistic sentence seems preferable, because arguments are not the kind of item which is true or false. Therefore I shall consider no condensed argument conditionals but rather meta-inference conditionals, where I stipulate that 'If p, then q' is a meta-inference conditional if and only if it is synonymous with ' 'q' can be inferred from 'p' '. I shall be considering meta-inference conditionals based on *non-deductive inferences*. In this respect my account differs from Nagel's.[18]

(2) Supposition/Assertion Conditionals and Rationality-judgment Conditionals.

In *Truth, Probability and Paradox,* Mackie proposes a more general account of conditionals than the condensed argument account. His account is based on suppositions. Thus he says "I want to offer, then, this general analysis: to say 'If P, Q' is to assert Q within the scope of the supposition that P."[19] The supposer gives his supposed situation or world some features including various features of the actual world he knows nothing about (for example, the supposer says 'Suppose the circumstances were like the actual ones except that Cleopatra's nose was longer'). He then can make discoveries about his supposed world. I now stipulate that 'If p, then q' is a *supposition/assertion conditional* if it is synonymous with 'Suppose p, then q'. I shall also be concerned with a meta-linguistic variant of supposition/assertion conditionals.

Roughly speaking, I stipulate that 'If p then q' is a

18. See Nagel, E., *The Structure of Science,* p.72.

19. Mackie, J.L., *Truth, Probability and Paradox,* p.93.

rationality-judgment conditional if it is synonymous with 'It is rational to believe q on evidence p'. The situation which I envisage is one in which a rational man who has a set of antecedent beliefs, B, is asked to suppose new evidence p (which might be inconsistent with his antecendent beliefs) and then he rationally asserts q. Thus I stipulate that 'If p, then q (relative to B)' is a rationality-judgment conditional if it is synonymous with 'Someone with set of antecedent beliefs B on supposing p could rationally assert q'.

Note One: If one is an objectivist about rationality-judgments then rationality-judgment conditionals would have truth-values. Even if one is not an objectivist about rationality-judgments, one would say that a rationality-judgment conditional is "true" or "false", where the scare-quotes indicate that the rationality-judgment conditional is usually, in pre-philosophical contexts, taken to be true or false. Since the sentence 'X causes Y' is also usually taken to be true or false, it seems that rationality-judgment conditionals are more suited to the analysis of causation than supposition/assertion conditionals.

Note Two: Of special importance is the case in which B, the set of antecedent beliefs, contains beliefs about laws of Nature. If a sentence p is *inconsistent* with these laws of Nature, the rational man is asked to suppose that there is an exception to some law of Nature. If it is objected that laws of Nature have no exceptions, then I suggest that the rational man supposes that some generalisation is not *universal* but is not *accidental* either.[20] For example, if the set

20. See pp.110-3 for a discussion of non-accidental generalisations.

of antecedent beliefs contains Newton's laws and if the rational man were asked to suppose that a thousand years ago a stone - for no reason - stayed two minutes above the earth without falling, then he would nonetheless assert that if, now, a stone is thrown up it will fall to the earth. Thus Newton's laws are supposed to be neither universal nor accidental.

Note Three: One respect in which meta-inference conditionals differ from rationality-judgment conditionals is the following:

If p is inconsistent with the set of antecedent beliefs B then one could, according to many accounts of deduction, infer *anything* from the set of beliefs $B \cup \{p\}$. However one might rationally *modify* B in the light of p and so obtain a *new* set of beliefs from which q can, but ~q cannot, be inferred. Thus there is, in this case, an important difference between the meta-inference conditional 'If p and B, then q' and the rationality-judgment conditional 'A person, with antecedent beliefs B, on supposing p could rationally assert q'.

Note Four: In Chapter Two, I shall consider meta-inference conditionals based only on certain specified rules of inference. I claim that *if* the new evidence p is consistent with the set of antecedent beliefs B *and if* the judgments of rationality were based only on the specified rules of inference, *then* the meta-inference conditional and the rationality-judgment conditional would be *equivalent*. Both would be analysed as:

'q' can be inferred from 'p' and the set of antecedent beliefs B using only the specified rules of inference.

Note Five: I assume that the supposed new evidence p itself belongs to the set of beliefs obtained when B is modified by p, that is, I suppose that one does not refuse to accept p. Also, if there are beliefs which one *refuses* to modify, these beliefs could conveniently be included with the new evidence. These assumptions are for the sake of simplicity. One could, instead, consider not simply the antecedent beliefs but the *strengths* of the various antecedent beliefs and consider not merely new evidence but the *strength* of the new evidence. The both antecedent beliefs and new evidence of low strength would, if necessary, be modified to preserve consistency.

(3) Lewis-conditionals

Both Stalnaker and Lewis have proposed accounts of conditionals based on possible worlds; I shall consider Lewis' account only. One assumes that there is a relation of comparative similarity between possible worlds. Lewis suggests that the conditional 'If it were the case that ϕ, then it would be the case that ψ 'is true at some world i if and only if *either*

(1) there is no world accessible from i at which ϕ is true; *or*

(2) there is a world k accessible from i at which ϕ is true and if some world j is at least as similar to i as k is, then $\phi \supset \psi$ holds at j.[21]

If one is only concerned with truth in the actual world and also one ignores the case in which the antecedent is *impossible*, Lewis' account is equivalent to:

---

21. Lewis, D., *Counterfactuals*, p.49.

> 'If p then q' is to be analysed as: 'There is a possible world w_1 in which p and q are both true, and if w_2 is any world in which p is true but q is false w_2 is less like the actual world than w_1 is.'

I shall stipulate that if 'If p then q' is a Lewis conditional it is synonymous with the above proposed analysans of 'If p, then q'.

CHAPTER TWO

CAUSATION AND CONDITIONALS

Part One of this chapter is principally a defence of a conditional analysis of causation based on Mackie's[1], but I also compare Mackie's account with Lewis'.[2] In Sections One and Two I expound Mackie's and Lewis' accounts without detailed discussion. In Section Three I consider the *Necessity Thesis*, namely that the analysis of 'X causes Y' is:

X and Y both occur and are distinct existences
and X is *necessary*-in-the-circumstances for Y.

I defend this thesis against some *preliminary* objections and, as part of this defence, I provide an account of the circumstances referred to. Finally in this section I argue that Lewis' account is unsatisfactory and the modifications required to make it satisfactory also make it converge with Mackie's account. Section Four is a discussion of the claim that, in addition to necessity, causes should be *sufficient* in the circumstances for the effect. I conclude that it is neither a necessary nor a sufficient condition for X to cause Y, that:

X and Y occur and are distinct existences
and X is *sufficient*-in-the-circumstances for Y.

In Sections Five, Six and Seven, I examine some further objections to the Necessity Thesis based on the problems of *Collateral Effects*, of *Underdetermination and Overdetermination*, and of *Simultaneous Causation*.

I modify the conditional analysis to meet these objections, obtaining the final version (*Nec 5*) in Section Seven.[3] Throughout

---

1. Mackie, J.L., *The Cement of the Universe*, Chapter 2, pp.29-58.
2. Lewis, D., "Causation", *Journal of Philosophy*, vol.70, (Oct.1973), pp. 556-569.
3. See p. 105.

Part One I use the Analysis-criterion,[4] and I ignore the possibility of backwards causation.

In Part Two I discuss what account might be given of the conditionals occurring in the analysis of causation. In Section Eight I discuss the notion of a non-accidental generalisation. In Section Nine I argue in favour of treating the conditionals occurring in the analysis of causation as (non-deductive) *meta-inference* conditionals: 'If p then q' is analysed as ' 'q' can be *inferred* from 'p' *and* the correct non-accidental generalisations'. In Section Ten I use the conditional analysis together with this account of conditionals to provide an account of the features on which causation is *supervenient*.

Section One: Mackie's Account of Causation in Terms of Conditionals

1.1 In Chapter Two of *The Cement of the Universe*, Mackie says:

> My main concern in this chapter has been to analyse our actual concept of causing, to formulate what we commonly take to be the distinguishing mark of causal sequences.[5]

Mackie's method is to begin by comparing a causal with an otherwise similar non-causal sequence, and to see what distinguishes the two. He compares:

> A: A chestnut is stationary on a flat stone. I swing a hammer down so that it strikes the chestnut directly from above. The chestnut becomes distinctly flatter than before.
>
> B: A chestnut is stationary on a hot sheet of iron. I swing a hammer down so that it strikes the chestnut directly from above. At the very instant that the hammer touches it, the chestnut explodes with a loud pop and its fragments are scattered around.[6]

---

4. See pp. 16-7.
5. Mackie, J.L., *The Cement of the Universe*, p.57.
6. *Ibid.*, p.29.

Mackie suggests that in A a counterfactual conditional holds, namely 'If the hammer had not struck the chestnut, the chestnut would not have become flatter'. But in B the analogous counterfactual does not hold; the chestnut would have exploded even if the hammer had not struck it.

Having obtained this initial suggestion for an analysis of causation, Mackie then uses the *method of counter-examples* as discussed in Chapter One, Section One of this thesis. On the whole, Mackie's examples are *empirically possible* ones. For even Mackie's indeterministic chocolate machines L and M could, perhaps, be constructed using a radio-active substance and a Geiger-counter.[7] The one exception occurs when Mackie says "We can coherently consider the possibility of backwards causation".[8] Mackie is there modifying an earlier account in order to accommodate cases of backwards causation, not on the grounds that backwards causation is empirically possible, but on the grounds that backwards causation is *coherent*. Thus Mackie seems to be using the *Analysis-criterion*.[9]

1.2 Mackie's Discussion of the Necessity Thesis

Mackie's initial suggestion is that 'X caused Y' means 'X occurred and Y occurred and Y would not have occurred if X had not.[10]

Mackie then considers a counter-example.(Because he denies that counterfactuals, in general, have truth-values, he does not present his counter-examples by arguing that analysandum and analysans have different truth-values. Rather, Mackie talks in terms of what one says and what one admits. However, I shall misrepresent Mackie on

---

7. Mackie, J.L., *The Cement of the Universe*, pp. 40-42.
8. *Ibid.*, p.52.
9. *Ibid.*, pp.16-7.
10. *Ibid.*, p.31.

this point and talk of the truth-values of analysandum and analysans. The words 'true' and 'false' could be put in scare-quotes.) The example is: Striking a match caused the appearance of the flame, but if the match had been touched with a red-hot poker then the flame would still have appeared. So, Mackie argues, the counterfactual 'If the match had not been struck the flame would not have appeared' is false. Thus the analysandum is true, the proposed analysans is false, so the analysis is incorrect. This counter-example is effective only if one gives some extra information, such as the information that someone intends to make the flame appear using, if necessary, a red-hot poker. For the counterfactual 'If the match had not been struck the flame would not have appeared' might be considered true even if there were other possible, but unlikely, sufficient conditions for the flame to appear. Mackie then modifies his analysis to:

> 'X caused Y' means 'X occurred and Y occurred and in the circumstances Y would not have occurred if X had not.'[11]

The phrase 'in the circumstances' is somewhat vague. One of my aims in discussing Mackie's account is to give a more precise account of the circumstances.

Also, Mackie points out that cause and effect are to be *distinct existences*. He notes that the description of cause and effect need not be logically independent, but he does not give an account of what is meant by 'distinct existences'.

Mackie makes a further modification to his analysis in order to handle examples of *collateral effects*. He gives the example in which Labour's defeat at the election pleases James but saddens John. In this example, James' being pleased does not cause John's being

---

11. Mackie, J.L., *The Cement of the Universe*, p.31.

sad, but one might say that *in the circumstances* John would not have been sad if James had not been pleased.[12] To handle these examples Mackie introduces the notion of *causal priority*; the earlier of two collateral effects is not causally prior to the latter. Mackie then proposes an anlysis of causal priority and so obtains a modified analysis of 'X caused Y', namely:

> X and Y occurred and if X were kept out of the world in the circumstances referred to and the world ran on from there, Y would not occur.[13]

At this point, Mackie further modifies his analysis in order to handle cases of backwards causation. For the above analysis is such that causes would have to precede effects. In Chapter Two, Mackie only sketches the modification required; one indirectly keeps X out of the world by keeping some event W, earlier than both X and Y, out of the world, and then one lets the world run on from there.

Mackie defends his analysis of causation against three other objections. The first objection is based on a proposed counter-example in which the analysandum is 'Jane's eating fishpaste in the pantry caused her feeling ill' and the analysans is 'Jane ate fishpaste in the pantry and felt ill and in the circumstances she wouldn't have felt ill if she hadn't eaten fishpaste in the pantry'.[14] Here it might seem that the analysandum is true but that the analysans is false. Mackie's reply is that if 'Jane's eating fishpaste in the pantry' is treated as a reference to an *event* (the same event as that referred to by 'Jane's eating fishpaste'), then both analysandum and analysans are true (provided 'Jane ate fishpaste in the pantry' is treated as asserting that the event occurred which is referred to

---

12. Mackie, J.L., *The Cement of the Universe*, p.33.
13. *Ibid.*, p.51.
14. *Ibid.*, p.33.

by 'Jane's eating fishpaste in the pantry'). But if the phrase 'Jane's eating fishpaste in the pantry' is treated as synonymous with 'the *fact* that Jane ate fishpaste in the pantry', then both analysandum and analysans are false. So in neither case is the analysis itself threatened. Now I have assumed that, in cases of *producing causation*, the cause, whether it is the occurrence of an event or the presence of a condition, is a *quasi-event*. Therefore, in Section Three I shall *restate* Mackie's reply to the objection in terms of whether the phrase 'in the pantry' is considered part of the *description* of the quasi-event or not.

The second objection is that if, for example, a block of flats explodes, one would not say that Jones' striking a match to light his cigarette was the cause of the explosion because it is assumed that people normally light matches in blocks of flats. In this case the analysandum 'Jones' striking a match to light his cigarette was a cause of the explosion' is false, but the proposed analysans 'Jones struck the match and the block of flats exploded and in the circumstances if Jones had not struck the match the block of flats would not have exploded' is true. To this Mackie replies that cause and effect are seen as *differences within a causal field* and anything that is part of the description of the field itself will be *automatically* ruled out as a candidate for the role of cause.[15]

Finally, Mackie considers at some length the problem of overdetermination. Mackie's account is satisfactory in cases of symmetric overdetermination provided one denies that either alternative is the cause.[16] He considers among others the following example:

---

15. Mackie, J.L., *The Cement of the Universe*, p.35.

16. Throughout this thesis I interpret examples of symmetric overdetermination "charitably". See the Note at the end of this section.

> Lightning strikes a barn in which straw is stored, and a tramp throws a burning cigarette butt into the straw at the same place and at the same time: the straw catches fire.[17]

One does not say, according to Mackie, that the lightning caused the fire. Now although the lightning was *sufficient* in the circumstances, it was not *necessary* in the circumstances; that is, the counterfactual 'In the circumstances, if the lightning had not struck the straw would not have burnt' is false.

However, cases of asymmetric overdetermination raise difficulties. Mackie considers among others the following ingenious example:

> A man sets out on a trip across the desert. He has two enemies. One of them puts a deadly poison in his reserve can of drinking water. The other (not knowing this) makes a hole in the bottom of the can. The poisoned water all leaks out before the traveller needs to resort to this reserve can; the traveller dies of thirst.[18]

In this example the analysandum 'Puncturing the can caused the death of the traveller' is considered true by Mackie, but the proposed analysans 'The can was punctured, the traveller died and in the circumstances if the can had not been punctured the traveller would not have died' is false. In this case Mackie argues that if the effect is treated as an *event* it could be described as *the death of the traveller by thirst*. In that case the proposed analysans *is* true. Mackie also claims that if the effect is treated as a *fact*, namely *the fact that the traveller died*, then the analysandum, namely 'Puncturing the can caused the fact that the traveller died' is false.

---

17. Mackie, J.L., *The Cement of the Universe*, p.44.

18. *Ibid.*, p.44.

Mackie also discusses whether what counts as a cause depends on the speaker's *interests*. He gives an example in which "an electric current in a certain wire (A), decayed insulation at a point on that wire (B), and inflammable material near that point (C)" are individually necessary and jointly sufficient for a house catching fire.[19] If one presupposes a set of circumstances one might treat, say, B and C as *the* circumstances, so that A but neither B nor C caused the fire. But if in the analysis of 'X caused Y' one is asserting that there are circumstances in which if *X had not occurred* Y would not have occurred, then A, B and C could all be causes of the fire. Mackie suggests that the presupposition account is better for catching the force of a causal statement *in use*, but the assertion account represents a desirable tidying up of the *meaning* of causal statements. It is worth noticing that although Mackie is using something like an analysis-criterion he claims that, strictly speaking, the analysandum and analysans have no truth-values. Hence Mackie resorts at this point to considerations of the *use* of sentences rather than their *truth-conditions*. If one ignores Mackie's scruples about the truth-values of counterfactuals, one could argue that his assertion account is a permissible *analysis* on the grounds that even if one were *interested* in the cause A, one would *admit* that B and C were *also* causes.

Note: In a case of symmetric overdetermination of C by A and B, I assume that if A but not B, or if B but not A occurred, then C occurs - not some slightly different quasi-event. Thus if two bullets (bullet number one and bullet number two) enter a man's heart simultaneously, I assume for the sake of the example that all the later events leading up to death are exactly as if only one bullet had

19. *Ibid.*, p.37.

entered the heart. In actual cases, if one considered enough detail, one would argue that both bullets were necessary in the circumstances for some event D which, in turn, causes death. Hence, even on the Necessity Thesis, the entry of bullet number one and the entry of bullet number two would be causes of D and consequently causes of death. The interesting case - but one in which people's intuitions disagree - is the *hypothetical* one in which the effect of A and B is *exactly the same* as the effect of A or B *alone*.

1.3 The Sufficiency Thesis

Mackie also discusses the plausible suggestion that part of the meaning of 'X caused Y' is that 'X is *sufficient* in the circumstances for Y'. As Mackie points out, if X and Y both occur then the conditional 'If X occurred, Y did' is automatically true. So Mackie proposes to discuss a strong conditional sense of 'sufficient-in-the-circumstances', namely 'In the circumstances, if Y had not been going to occur then X would not have occurred'[20]; that is, the *absence* of Y is *necessary*-in-the-circumstances for the *absence* of X. Mackie argues that although, in general, causes are sufficient-in-the-circumstances for their effects, there are counter-examples to the Sufficiency Thesis. He considers three chocolate bar machines K, L and M. K is a deterministic machine. But neither L nor M are deterministic; their operation involves random processes.[21] A shilling in the slot is necessary but not sufficient for L to produce a chocolate bar, and a shilling in the slot is sufficient but not necessary for M to produce a chocolate bar. Mackie suggests that one would say that putting the shilling into machine L caused the bar to be produced,

20. Mackie, J.L., *The Cement of the Universe*, p.39.
21. *Ibid.*, pp.40-42.

but putting the shilling into machine M was not a cause of producing the chocolate bar, even if in both cases a bar of chocolate was produced. If Mackie is correct here, then he has shown that the Sufficiency Thesis can be rejected. Mackie supports his claim by arguing that one is prepared to say 'Tom's saying what he did caused Bill to hit Tom' without prejudging whether human beings are deterministic.[22] He gives other proposed counter-examples to the Sufficiency Thesis:

> There being a radon atom here now is a causal consequence of there having been a radium atom here a little earlier; [23]

and

> Tom's being colour-blind is a causal consequence of Jim's being colour-blind.[24]

In both cases it is plausible that the processes are non-deterministic.

However, Mackie also provides arguments in favour of the Sufficiency Thesis. He argues that if P caused Q one can say that Q occurred because P did. He then claims that 'Q occurred because P did' is practically equivalent to Goodman's *factual conditional*, which is a combination of the conditional 'If P occurred Q occurred' with the presupposition that P did occur.[25] Mackie also argues that the future causal statements 'P will cause Q' and 'P would cause Q' seem to claim that P is sufficient in the circumstances for Q. Although I shall accept something like Mackie's conclusion, I find both these arguments confusing because they seem only to show that causes are *weakly* sufficient for their effects. Consider the case of the future and conditional causal sentences 'P will cause Q' and 'P would cause Q'; although weak sufficiency is entailed by the occurrence of P and Q, it is nonetheless important that P is weakly

---

22. Mackie, J.L., *The Cement of the Universe*, p.43.
23. *Ibid.*, p.48. 24. *Ibid.*, p.48. 25. *Ibid.*, pp.48-49.

sufficient for Q; that is, the conditional 'If P occurs, Q occurs' holds. For if one believes that P will occur and believes also that if P occurs Q occurs, then *on those grounds* one believes that Q occurs and hence that P will cause Q. Again, if one speculates that P might occur, and believes that if P occurs Q occurs, then one believes that P would cause Q. Thus I cannot understand how Mackie's arguments are intended to establish a *strong* sufficiency thesis based not on the conditional 'If P occurs Q occurs', but on the *non-equivalent* conditional 'If Q were not going to occur, P would not occur'.

1.4 Mackie's Account of the Origins of the Concept of Causation

Because Mackie holds that counterfactuals are supposition/assertion conditionals the counterfactuals used in the analysis of causation have, in general, no truth-values. Furthermore, even if it is claimed that counterfactuals have truth-values, their truth cannot be observed in the actual sequence which is described as causal. Hence Mackie considers that a *psychological* explanation should be given for why people assert some counterfactuals rather than others. On Mackie's account of counterfactuals one requires an explanation for how people decide what to assert within the scope of a supposition. Mackie suggests that the *sophisticated* way of deciding what to assert is a combination of inductive and deductive reasoning. Presumably one uses what one believes to be non-accidental (but perhaps not universal) generalisations, together with what is supposed (which might, perhaps, be an exception to the corresponding universal generalisation), and one makes the assertion accordingly. The *primitive* method according to Mackie is based on imagination and analogy. Finally, he considers the concept of causal priority, the

origins of which he sees, plausibly enough, in the experience of intervention in the world. If one introduces a change X and change Y occurs, then one sees X as *causally prior* to Y.

Section Two: A Brief Account of Lewis' Theory of Causation

Lewis' article "Causation"[26] is not intended as a completely satisfactory account of causation. In particular, Lewis confines himself to causation among events, and to an analysis "that works properly under determinism". Nonetheless, his account is important, and I shall want both to criticise it and to borrow from it certain features when I modify Mackie's account. As the two accounts are modified they converge.

Lewis defines causal dependence as follows:

> If c and e are events and if O(c) and O(e) are the propositions that c occurs and e occurs respectively, then e is causally dependent on c if the two Lewis-conditionals 'If O(c) then O(e)' and 'If ~O(c) then ~O(e)' are true.

As Lewis points out, the conjunction of O(c) and O(e) entails the Lewis-conditional 'If O(c) then O(e)', which thus corresponds to what Mackie calls *weak* sufficiency. Hence Lewis' account is essentially a *Necessity Thesis*.

Next, Lewis analyses 'c causes e' as:

> c and e occur and there is a finite sequence of events of which c is the first, of which e is the last and such that each member (except the first) is causally dependent on the previous member.

---

26. Lewis, D., "Causation", *Journal of Philosophy*, Vol.70 (Oct.1973), pp. 556-569.

Lewis requires these causal chains to ensure that causation is *transitive*. For the inference-pattern

'If ϕ then ψ;
if ψ then χ
$\therefore$ If ϕ then χ'

is not valid for Lewis-conditionals (or for counterfactuals in general).[27] However, it is worth noticing that if one also requires what Mackie calls a *strong* sufficiency thesis, in this case the Lewis-conditional 'If ~O(e) then ~O(c)', then one would not require causal chains. For the inference pattern

'If ϕ then ψ
if ψ then ϕ
if ψ then χ
$\therefore$ if ϕ then χ'

is valid for Lewis-conditionals.[28]

In "Causes and Counterfactuals", Kim[29] argues that the "sort of dependency expressed by counterfactuals is considerably broader than strict causal dependency": that is, Kim argues that Lewis' account provides at best sufficient but not necessary truth-conditions for causation. Kim gives four kinds of example to support his claim.

(1) There are cases in which there is a logical not a causal connection, for example 'If yesterday had not been Monday, today would not be Tuesday.'

(2) There are cases in which one event is part of another. For example 'If I had not written 'r' twice in succession, I would not have written 'Larry'.'

(3) There are cases involving actions of agents. For example 'If I had not turned the knob, I would not have opened the window.'

(4) There are examples like 'If my sister had not given birth at t, I would not have become an uncle at t.'

---

27. Lewis, D., *Counterfactuals*, p.32.

28. *Ibid.*, p.33.

29. Kim, J., "Causation and Counterfactuals", *Journal of Philosophy*, vol.70 (Oct.1973), pp.570-572.

I claim that the first three examples are covered by some requirement that cause and effect be distinct existences. In the case of example (3), the action of turning the knob is contained in the action of opening the window. Example (4) is similar to the example of Joanne giving birth resulting in Harry becoming a father,[30] which is also ruled out as a case of causation by a less straightforward distinct existences requirement.

Thus I suggest that what is required to make Lewis' account initially plausible is merely a *distinct existences* requirement.

Section Three: The Necessity Thesis

In this section my aim is to defend against some *preliminary* objections a version of the Necessity Thesis (*Nec 2*), based on Mackie's account. Also, I shall begin my criticism of Lewis' account of causation.

3.1 A preliminary version of the Necessity Thesis as stated by Mackie, is:

Nec 1: 'X caused Y' is analysed as

(1) X and Y are distinct existences and both occurred.

and

(2) In the circumstances Y would not have occured if X had not.

That the truth of the proposed analysans is, in all coherent examples, necessary for the truth of the proposed analysandum I shall call the *necessity* of the Necessity Thesis; that the truth of the analysans is sufficient I shall call the *sufficiency* of the necessity Thesis. I shall now discuss the initial objections which Mackie considers to the Necessity Thesis.

---

30. See p. 44.

Objection One (To the necessity of the Necessity Thesis)

Mackie discusses the example 'Jane's eating fishpaste in the pantry caused her feeling ill'[31] and the suggested expansion 'Jane ate fishpaste in the pantry and felt ill and in the circumstances she wouldn't have felt ill if she hadn't eaten fishpaste in the pantry'. The objection is that the analysandum is true but the analysans is false; if Jane had eaten the fishpaste in the kitchen she would still have been sick.

Reply: Mackie claims that either one would reject both the analysans and the analysandum, or one would accept them both. Mackie's claim is based on a separation of the case in which the cause in an *event* from the case in which the cause is a *fact*. However, I am considering causes and effects to be quasi-events, so I restate Mackie's argument as follows:

There are at least two possible paraphrases for the sentence 'Jane's eating fishpaste in the pantry caused her feeling ill':

Paraphrase One: The quasi-event described by 'Jane ate fishpaste *in the pantry*' caused the quasi-event described by 'Jane felt ill'.

Paraphrase Two: The quasi-event described by 'Jane ate fishpaste' caused the quasi-event described by 'Jane felt ill', *and* Jane ate the fishpaste in the pantry.

The difference in spoken English would be one of stress. If the phrase 'in the pantry' were stressed, then Paraphrase One would be appropriate. Otherwise, Paraphrase Two would be more appropriate. Now, if the place where the fishpaste is eaten is irrelevant, Paraphrase One is false and Paraphrase Two is true. The analysans for

---

31. Mackie, J.L., *The Cement of the Universe*, p.33.

Paraphrase Two would also be true. So in this case the difficulty - such as it is - is in paraphrasing the ordinary English causal sentence into a sentence about the causal relation between quasi-events; the analysis of sentences about the causal relation between quasi-events is not threatened.

There is an objection to this reply which should be met. The objection is that someone could correctly utter 'Jane's eating the open jar of fishpaste caused her feeling ill' *without committing himself* to whether the fact that the jar of fishpaste was open was causally *relevant or not*. Hence it might be claimed that neither of the above paraphrases is suitable, and that one should adopt a sufficiency thesis rather than a necessity thesis. Such cases I consider to be sentences involving causal overdetermination due to the ignorance of the speaker who does not know what is the necessary condition, but who merely indicates an area in which a necessary condition is to be found.

I claim that one need not abandon the necessity of the Necessity Thesis, for one can retain the Necessity Thesis when discussing causal relations between quasi-events provided suitable care is taken in paraphrasing the ordinary English into sentences about the causal relations between quasi-events. Thus in the case of 'Jane's eating the open jar of fishpaste caused her feeling ill' the paraphrase is the following:

Either the quasi-event described by 'Jane ate the *open jar* of fishpaste' caused the quasi-event described by 'Jane felt ill';

or the quasi-event described by 'Jane ate a jar of fishpaste' caused the quasi-event described by 'Jane felt ill', *and the jar which Jane ate was open.*

Again, suppose one says 'The weather caused sickness in the sheep'. This could be paraphrased as:

> There was some deviation from normal weather conditions, which caused the sickness in the sheep.

Another preliminary objection to the Necessity Thesis is:

Objection Two: (To the sufficiency of the Necessity Thesis) Some conditions are not causes but are nonetheless *necessary* for the effect. For example, the presence of oxygen is necessary for the fire. Hence the Necessity Thesis is not, it might be claimed, sufficient for X to cause Y.

Reply: I have already dealt with this kind of objection in Chapter One, Section Two.[32] The presence of oxygen is not a quasi-event because all normal states of affairs are ones where oxygen is present. This is essentially the same reply as Mackie's.[33]

A third objection which might be made to Mackie's Necessity Thesis is:

Objection Three: The phrase 'in the circumstances' is rather vague. If there are several jointly sufficient and individually necessary causes A, B etc. for E, is the occurrence of B etc. to be considered part of the circumstances or not? Are the events such as the presence of oxygen part of the circumstances? What is *not* part of the circumstances?

Reply: I admit that the phrase 'in the circumstances' is rather vague and should be made more precise. If one is discussing a necessity thesis the phrase 'in the circumstances' must be interpreted as ruling out sufficient conditions that did not in fact occur. Thus if one claims that the faulty wiring was necessary in the circumstances

---

32. See pp. 31-35.

33. Mackie, J.L., *The Cement of the Universe*, p.35.

for the house to burn, one is ruling out sufficient conditions which did not occur. For example, the faulty wiring was not absolutely necessary for the house to burn; it might have been struck by lightning. Now it might be argued that the counterfactual 'If there had not been faulty wiring the house would not have burnt' is true, although some extraordinary event such as a lightning strike would have been sufficient for the house to burn. Indeed, even if the sufficient condition that did not occur was not at all extraordinary, the counterfactual might be true. Suppose the lightning strike caused the house to burn but it had only been exceptional good luck that the inflammable material was never near the fault in the wiring. In that case one would still say 'If the lightning had not struck then the house would not have burnt'. This might seem to show that the phrase 'in the circumstances' is redundant. However, there are two reasons for not simply eliminating the phrase 'in the circumstances'. The *first* is that ordinary counterfactuals without a description of the circumstances are as vague as the phrase 'in the circumstances' itself. When I modify the Necessity Thesis to handle, say, cases of collateral effects, I require the precision obtained by explicitly mentioning the circumstances in the conditional. The *second* reason is that, while I claim that the truth of the meta-inference conditionals or rationality-judgment conditionals is sufficient for the truth of corresponding ordinary counterfactuals, I *do not claim* that *all counterfactuals* correspond to suitable *meta-inference* or *rationality-judgment* conditionals. It is by *explicitly mentioning the circumstances* that one obtains counterfactuals which correspond to meta-inference or rationality-judgment conditionals. Thus when discussing the Necessity Thesis one should mention the circumstances, which involves mentioning all the quasi-events that do not occur but would

be (causally) sufficient if they were to occur. To avoid the circularity due to the phrase 'causally sufficient', I shall stipulate that the circumstances (at the time of the cause) are such that *no* quasi-events occurred that did not in fact occur.

Now one might also be interested in the notion of *sufficiency-in-the-circumstances*. If, which is plausible, the faulty wiring is to be sufficient-in-the-circumstances for the house to burn then, obviously, the other necessary conditions which actually occurred are to be part of the circumstances. Again, to avoid circularity, one supposes that all the quasi-events occurred which actually occurred (about the time of the cause) and which were distinct existences from the proposed cause X. The proviso that the circumstances consist of quasi-events which are *distinct existences from X* is required in order to prevent the circumstances, Z, containing the occurrence of X itself and so to prevent one obtaining the conditional 'If Z occurred and X had not occurred then Y would not have occurred'. Surely inconsistent antecedents in conditionals should be avoided.

For the moment I shall assume that one need only consider quasi-events that occur no earlier than and no later than X; that is, the quasi-events occupy time-intervals contained in that occupied by X. In particular, if X causes Y which causes Z, then X causes Z - at least in the vast majority of cases - but if one were to suppose *as part of the circumstances*, that Y occurred then X would not be necessary in the circumstances for Z.

In order to make the Necessity Thesis more precise I *stipulate* that the conditional:

> In the circumstances <occurring in T; distinct existences from Q_1 ... Q_m> if p then q

is to be interpreted as the conditional 'If p & r & s then q' where:

r is 'R_1 ..., R_n occur' and R_1 ... R_n are the quasi-events which actually occurred in the time-interval T and which were distinct existences from Q_1 ... Q_m;

and where:

s is 'No quasi-event occurred in T other than those contained in R_1 & ... & R_n & Q_1 & ... & Q_m'

Thus the circumstances in the interval T are supposed to be exactly as they actually were, except that it is not specified whether Q_1 ... Q_m occurred or not. Accordingly, I propose a more precise version of the Necessity Thesis, namely *Nec 2:* 'X causes Y' is to be analysed as:

(1) X and Y are distinct existences and both occur

and

(2) In the circumstances <occurring no earlier or later than X; distinct existences from X> if X does not occur Y does not occur.

Note 1: *Nec 2* is similar to Lyon's *condition 1.*[34]

Note 2: I shall modify *Nec 2* in subsequent sections until I eventually obtain *Nec 5.*[35]

Note 3: For convenience, I use the tenseless unmodalised form of the conditional.

Note 4: Any state of affairs which is part of all the normal states-of-affairs and which occurs in T is also assumed to be part of the circumstances.

---

34. Lyon, A., "Causality", *British Journal for the Philosophy of Science,* vol.18 (1967), p.8.

35. See p. 105.

Note 5: If two quasi-events A_1 and A_2 symmetrically overdetermine B,[36] as in Mackie's example of the lightning striking the barn at the same time as the tramp throws down a burning cigarette butt, neither A_1 nor A_2 is necessary in the circumstances for B. This *initially* seems to support the Necessity Thesis, since one is reluctant to call either A_1 or A_2 a cause. However, one is also reluctant to deny that A_1 and A_2 are causes of B. Consequently, such examples of symmetric overdetermination present some difficulty for the Necessity Thesis.

3.2 I shall now make the first step in my criticism of Lewis' account of causation.

Consider the following hypothetical example. A new isotope of strontium is discovered with the remarkable property that if a nucleus of this isotope is simultaneously hit with three or four neutrons it emits an α-ray. Otherwise it does not. It is very rare that the nucleus is hit at all. On one occasion *three* neutrons hit the nucleus. I suggest that one would say that hitting the nucleus with three neutrons caused the α-ray to be emitted. Now in the supposed circumstances, if it is not the case that the nucleus is hit with three neutrons, then presumably the nucleus is hit by at most two neutrons, since it is rare for the nucleus to be hit at all. Hence no α-ray is emitted. However, on the obvious interpretation of Lewis' account it would seem that a world in which four neutrons hit the nucleus is as like the actual one as a world in which two neutrons hit the nucleus, and is more like the actual one than a world in which no neutrons hit the nucleus.

A defender of Lewis' analysis might argue that hitting the nucleus with three neutrons is the same *event* as hitting the nucleus with three or four neutrons. More plausibly, the defender of Lewis'

36. See Note on p. 61.

analysis might argue that hitting the nucleus with three neutrons is not a cause, since we have a case of *overdetermination* where the actual cause is hitting the nucleus with three or four neutrons. Perhaps Lewis' analysis can be defended in this way, but it seems to me that one would say 'Hitting the nucleus with three neutrons *caused* it to emit an α-ray'. I suggest, therefore, that in order to defend Lewis' account, one would have to claim that the world in which no neutrons hit the nucleus is more like the actual one than the world in which four neutrons hit the nucleus.

But now it seems that Lewis' account is rather imprecise. What are Lewis' rules for judging the similarity of worlds? There should be some rules; for surely the similarity of two worlds is *supervenient* on what happens in those two worlds. If not, the similarity of two worlds would seem to be contingent and so, presumably, there are *possible superworlds* corresponding to possible similarity relations between possible worlds. I shall argue later[37] that Lewis is committed to some rather peculiar rules for judging the similarity of worlds.

Section Four: The Sufficiency Thesis

Does admitting that X causes Y commit one to accepting that X was sufficient in the circumstances for Y? Mackie concludes:

> The general notion of a cause is of something which is both necessary and sufficient in the circumstances for its effect, but where the cause and the effect have both actually occurred we do not require that the cause should be sufficient ...[38]

In this section I shall examine some arguments, which I do not consider decisive, against the necessity of the Sufficiency Thesis. I shall also discuss the appropriate analysis of the Sufficiency Thesis.

---

37. See pp. 89-91.
38. Mackie, J.L., *The Cement of the Universe*, p.49.

4.1 I first examine the proposed counter-examples to the necessity of the Sufficiency Thesis.

Example One: Mackie's Chocolate Machine L

> The machine L contains a mechanism such that unless a shilling is put in no chocolate bar comes out, but if a shilling is put in, the chocolate bar coming out is a random event - say the chocolate bar comes out on fifty percent of occasions.[39]

The shilling being put in is not then sufficient in the circumstances for the chocolate bar to come out, for in *similar* circumstances the bar sometimes comes out yet sometimes does not. Now Mackie claims that if a shilling is put in and a chocolate bar comes out then the shilling being put in *caused* the chocolate bar to come out. If he is correct about this causal statement, then he has a strong argument against the necessity of the Sufficiency Thesis. However, I suggest that the claim that the shilling's going in caused the bar's coming out only seems plausible because one tends to confuse the chocolate machine L with another chocolate machine which is sometimes in a mysterious unobservable state S; the machine being in S, *together with* the shilling being put in *is* sufficient for the production of a chocolate bar, so that *in the circumstances* putting a shilling in is sufficient. My claim is supported to some extent by Mackie's account of chocolate machine M, which sometimes produces chocolate bars without the shilling being put in. Mackie assimilates this example to cases of causal overdetermination.[40] But surely this assimilation is only plausible if there is a mysterious unobservable state which, when it occurs, is by itself sufficient to produce a chocolate bar.

---

39. Mackie, J.L., *The Cement of the Universe*, pp. 41-3.
40. *Ibid.*, pp. 42-3.

Example Two: "Tom's saying what he did caused Bill to hit Tom." [41]

Mackie suggests that even if Bill is free not to hit Tom, the sentence might be considered true. But if Bill is free not to hit Tom, Tom's saying what he did is not sufficient in the circumstances for Bill to hit him. As in the previous example, one tends to assume that there is an unobservable state S - in this case a mental one - and that if S is part of the circumstances, Tom's saying what he did is sufficient for Bill to hit Tom. It is not obvious that someone could both deny the existence of the state S and assert that Tom's saying what he did caused Bill to hit him.

Example Three: Mackie says that "There being a radon atom here now is a causal consequence of there having been a radium atom here a little earlier."[42]

Would one say that the presence of the radium atom caused the presence of the radon atom? If so, then here is another counter-example to the *necessity* of the Sufficiency Thesis.

This example is less confusing since in this case one does *not* tend to assume the existence of unobservable states. If one asked the question 'What, if anything, did the presence of the radium nucleus cause?' one would not, I think, answer 'The presence of the radon nucleus'. But if one asked 'What, if anything, caused the presence of the radon nucleus?', one might reply 'The presence of the radium nucleus'. I am not *sure* how this difference should be interpreted. One possible interpretation is that there are two pre-philosophical notions; a cause considered as something necessary and a cause considered as something sufficient. In a context where it is known that X

41. Mackie, J.L., *The Cement of the Universe*, p.43.
42. *Ibid.*, p.48.

occurs and one is speculating about what other events occur (using one's beliefs about what can cause what) if one asks 'What will X cause?" or What did X cause?', one is concerned with causes which are (weakly) *sufficient* in the circumstances for their effects, since the truth of the answer 'Y' depends only on the (weak) *sufficiency* of X for Y and the occurrence of X. Conversely, if one asks 'What caused X?' and the answer is 'Z', the truth of the answer depends only on the occurrence of X and the *necessity* in the circumstances of Z for X, so one is interested in causes that are *necessary* in the circumstances for their effects. However in the context in which the person being asked 'What caused X?' had *observed* what happened but the person asking the question had not, the person being asked would tend to seek some event or condition which occurred and which was causally the most important feature; ideally it would be *both* necessary *and* (strongly) sufficient for X (in the circumstances), but otherwise it would be *either* necessary *or* sufficient for X. In the case of symmetric overdetermination there is no *single* causally most important feature, so one does not know whether to give all or none of the features as the cause.

There is another argument against the necessity of the Sufficiency Thesis, which is due to Martin. He claims that if the Sufficiency Thesis is accepted as a necessary condition for A to cause P, then evidence E supports the claim that A was the cause of P *only to the same degree* that it supports the claim that A was sufficient for P.[43] For example, suppose a chocolate bar machine otherwise like L is in fact deterministic. Although I do not know what the mechanism is, I believe there *is* a mechanism, and that an explanation

---

43. Martin, R., "The Sufficiency Thesis", *Philosophical Studies*, vol. 23 (1972), p.210.

could, in principle, be given for the machine's erratic behaviour. In that case I have little evidence that the shilling is sufficient for the chocolate to come out. Yet I would confidently assert that putting the shilling is was a cause of the chocolate coming out.

An objection to Martin's argument is that, until I have seen the chocolate come out, I would *not* confidently assert that the shilling going in will cause the chocolate to come out. But after I have seen the chocolate come out I could argue as follows:

> Since I believe the machine is deterministic, there is some sufficient condition for the chocolate bar to come out. The only noticeable change is that the shilling goes in. So, especially if I have used the machine on several occasions, I have good evidence that there was not the sufficient condition *before* I put the shilling in, that there was the sufficient condition *afterwards*, and hence that putting the shilling in is sufficient *in the circumstances* (of which I am ignorant).

The above discussion of the proposed arguments against the Sufficiency Thesis, and the uncertainty of one's intuitions in the (charitably interpreted) symmetric overdetermination case, show, I suggest, that the notion of a cause is more complicated than is claimed by adherents either of the Necessity Thesis or of the Sufficiency Thesis. I *suggest* that what is considered to be a cause depends on the context - not merely because the causal field depends on the context.

(1) If both the Necessity Thesis and the Sufficiency Thesis are satisfied, X causes Y.

(2) If neither the Necessity Thesis nor the Sufficiency Thesis

is satisfied, X does *not* cause Y.

(3) In contexts where an event X is known to occur and one *therefore* infers that some earlier event X also occurs, then for X to cause Y the (weak) Sufficiency Thesis is sufficient.

(4) In contexts where an event Y is known to occur and one *therefore* infers that some earlier event X also occurs, then for X to cause Y the Necessity Thesis is sufficient.

(5) In contexts where one asks an observer 'What caused Y?' or 'What did X cause?' when the observer had seen what had occurred then the observer, if he *cannot find* a quasi-event satisfying *both* the Necessity *and* the Sufficiency Thesis, may give as a cause a quasi-event which satisfies the Necessity Thesis alone *or* (perhaps) the Sufficiency Thesis alone.

Thus I do not decide between the two theses. I continue to modify both the Necessity Thesis and the Sufficiency Thesis, to meet various objections.

By analogy with *Nec 2*, I propose, as a *preliminary* version of the *Sufficiency Thesis*

Suff 1: 'X causes Y' is analysed as:

(1) X and Y occur and are distinct existences

and

(2) In circumstances <occurring no earlier or later than X; distinct existences from X> if Y does not occur X does not occur.

Note One: Modifications are made to *Suff 1* until I finally obtain *Suff 5*.[44]

Note Two: *Suff 1* is a version of the Sufficiency Thesis which involves strong sufficiency. Yet in the above discussion I mention

44. See Note Five on p. 108.

weak sufficiency. This apparent discrepancy is removed when I discuss probabilistic versions of the Sufficiency Thesis. *Suff 2*, which is similar in form to a weak sufficiency thesis, is stronger in one respect than *Suff 2**, which is similar in form to a strong sufficienty thesis. I argue in favour of *Suff 2*.[45]

Note Three: The Lewis-type analysis analogous to *Suff 1* would involve the Lewis-conditional 'If ~O(e) then ~O(c)'.

4.2 An objection to the analysis *Suff 1* is that, even in the circumstances, A might not *determine* B. Anscombe has argued that causes do not in general *determine* their effects.[46] Consider the example of the photographic plate which for two months is in a drawer next to a minute quantity of uranium (of such a size that there is a 99% chance that the uranium emits enough radiation in the two months to spoil the photographic plate). If the plate is spoilt one would say that the uranium spoilt it; in other words, the presence of the uranium *caused* the plate to become useless. The presence of the uranium is necessary in the circumstances but it is not absolutely sufficient in the circumstances; it is only *nearly* sufficient in the probabilistic sense that there is a high probability of the plate being spoilt. Therefore I suggest that *Suff 1* should be replaced by *Suff 2*: 'X causes Y' is analysed as:

(1) X and Y occur and are distinct existences

and

(2) In circumstances <occurring neither earlier nor later than X; distinct existences from X> if X occurs Y is likely to occur.

Note One: This is a *factual* conditional rather than a *counterfactual*

45. See p. 83.

46. Anscombe, G.E.M., "Causality and Determination", reprinted in *Causation and Conditionals*, ed. Sosa, E., pp. 63-81.

conditional. In the analysis of 'X caused Y' one would have 'In circumstances ... if X occurred [as it did] then Y was highly likely to occur'. This is not a trivial conditional even if one knows that Y occurs, since Y could have occurred even if it were highly improbable.

Note Two: I shall not consider what *kind* of probability is mentioned in the phrase 'Y is highly likely to occur' until I discuss the account to be given of the conditionals used in the analysis of causation.[47]

Likewise, one seems to require a *probabilistic* variant of the analysis in terms of Lewis-conditionals. Instead of the Lewis-conditional 'If the photographic plate were not going to become useless, the uranium would not be near it', one might have the variant 'There is some world w_1 in which the plate does not become useless and the uranium is not near it in the vast majority of those worlds which are at least as like the actual world as w_1 and in which the plate does not become useless'. Here w_1 has to be chosen so that there are *enough* worlds at least as like the actual world as w_1 for the phrase 'the vast majority' to be applicable.

Notice that the corresponding probabilistic variant of the Necessity Thesis does not provide an acceptable analysis of causation. Suppose there is a chocolate machine N which, unlike Mackie's machines L and M, is deterministic, but which is nonetheless erratic. When a shilling is put in N a chocolate bar comes out; very occasionally (depending on unobserved features such as the position of rusty cogs), N produces a *free* chocolate bar. Thus one could say that the probability of a shilling having been put in, given that a chocolate bar comes out, is very high. Nonetheless, one should not say that the

47. See Note Two on p. 123.

shilling *caused* the chocolate bar to come out but *rather*, that it is highly probable that the cause was the shilling being put in.

An alternative to *Suff 2* would be the following analysis *Suff 2** of 'X causes Y' as:

(1) X and Y occur and are distinct existences

and

(2) In circumstances <occurring no earlier or later than X; independent of X> if Y does not occur X is highly unlikely to have occurred.

Suff 2 concerns prediction; *Suff 2** concerns retrodiction. Now X and Y are quasi-events so they differ from some norm. Thus one can assume it is not the case that X and Y are very highly probable. In that case, if it is rational to predict that Y occurs on the evidence that X occurs, it would *also* be rational to *retrodict* that X does not occur on the evidence that Y does not occur. This is illustrated by the following argument involving probabilities:

By Baye's Rule

$\text{Prob}(X|\overline{Y}) \times \text{Prob}(\overline{Y}) = \text{Prob}(\overline{Y}|X) \times \text{Prob}(X)$

Prob(Y) is not too near 1 (say Prob(Y) is less than 90%)

$\therefore$ the ratio, $\text{Prob}(X) : \text{Prob}(\overline{Y})$ is not greater than 10 : 1.

$\therefore$ if $\text{Prob}(Y|X)$ is very high (say 99%), $\text{Prob}(\overline{Y}|X)$ is very low (1%) and so $\text{Prob}(X|\overline{Y})$ is fairly low (at most 10%)

$\therefore$ $\text{Prob}(\overline{X}|\overline{Y})$ is fairly high (at least 90%).

Hence, if the prediction is *very likely* to succeed, the retrodiction is *fairly likely* to succeed. The converse does not hold since if X is extremely unlikely anyway, it is safe to retrodict $\overline{X}$, regardless of what Y is. Since, in general, 'A causes B' licenses both prediction and retrodiction, I prefer the analysis *Suff 2* to *Suff 2**.

It also follows that even the probabilistic modification of the Lewis-type account of sufficiency is inadequate. Perhaps one could use the following Lewis-type analysis.

'c causes e' is analysed as:

(1) c and e occur and are distinct existences;

(2) There is a world w_1 in which O(c) and O(e) are both false;

(3) There is a world w_2 no more like the actual world than w_1, such that in w_2, O(c) and O(e) are both true;

and

(4) O(e) is true in the vast majority of worlds w such that
(i) w is at least as like the actual one as w_2
and
(ii) O(c) is true in w.

Note: Necessity, Sufficiency, and the Paradigms of Causation

The following two claims are initially plausible:

(1) If a causal situation in which X causes Y resembles those producing paradigms in which an agent *brings* X *about* in order to *bring* Y *about*, then the Sufficiency Thesis holds;

(2) If a causal situation in which X causes Y resembles those producing paradigms (the "preventing paradigms") in which an agent could have *prevented* Y by *preventing* X, then the Necessity Thesis holds.

But these claims represent an oversimplified view. Consider first Mackie's indeterministic chocolate machine L. A shilling is necessary but not sufficient in the circumstances for a chocolate bar. Nonetheless, if a shilling has been put in and a chocolate bar comes out, the situation closely resembles a "producing" producing paradigm; the shilling was put in *in order that* the chocolate bar should come out. Now consider the case of Mackie's chocolate machine M. A shilling is sufficient but not necessary in the circumstances for a chocolate bar.

In this case one does not have a "preventing" paradigm for one might not have succeeded in preventing the chocolate bar coming out by preventing the shilling going in. But an agent can ensure that a chocolate bar comes out by putting a shilling in, so the situation resembles a "producing" producing paradigm. Hence, an emphasis on the "producing" producing paradigms, in which an agent *produces* the cause in order to produce the effect, does not distinguish between the two theses. But if one requires resemblance to the "preventing" paradigms, then one has some grounds for requiring the necessity of the Necessity Thesis.

Section Five: The Problem of Collateral Effects

In this section I modify *Nec 2* to meet counter-examples based on collateral effect situations. I also modify *Suff 2*.

5.1 Mackie gives the following example of the problem of collateral effects.[48] Labour's defeat at the election pleases James but saddens John. Even assuming that John is saddened after James is pleased, James's being pleased does not cause John's being saddened. Yet one might claim that if James had not been pleased, John would not have been saddened. More generally, if C is a cause of A which is absolutely (not just probabilistically) sufficient in the circumstances for A and if C also causes B, then if A is earlier than B, A should, using *Nec 2*, cause B. This conclusion is supported by a valid argument for Lewis-conditionals.

If A had not occurred C would not have occured,
If C had not occurred A would not have occurred,
If C had not occurred B would not have occurred,
∴ if A had not occurred B would not have occurred.

This is an instance of the *valid* inference-pattern:

48. Mackie, J.L., *The Cement of the Universe*, p.33.

if ϕ then ψ ;
if ψ then ϕ ;
if ψ then χ ;
$\therefore$ if ϕ then χ .[49]

On the proposed version of Mackie's analysis (*Nec 2*), the problem does not arise in every case of collateral effects. In particular, suppose that all quasi-events occur which are distinct from A and which actually occur about the same time as A, and suppose that A does not occur. In that case C would *not* occur. But there would usually be an *intermediate* cause between C and B which is simultaneous with A and so is assumed to occur. Hence B would still occur. However, an analysis should be proof against conceivable *hypothetical* counter-examples. There are hypothetical examples which *reintroduce* the problem of collateral effects. For many people would claim that it is conceivable that, what Russell calls *mnemic causation*, should occur. That is, C causes B, there is a temporal gap between C and B, yet there are *no intermediary causes* connecting C and B in a causal chain. I do not want my defence of causal anti-realism to depend on whether mnemic causation is conceivable or not. *Here* I shall assume that mnemic causation is conceivable. *Elsewhere*[50] I shall show how to modify my account to *rule out* mnemic causation. So I assume that John learns of Labour's defeat at the same time as James, that John is not saddened at the time but becomes sad an hour later and that there is no chain of causes or mechanism of any kind - mental or physical - connecting John's learning the news and John's becoming sad; the learning of the news *directly* makes John sad an hour later. In that case one does not suppose the occurrence of any intermediary cause and so, if James had not been pleased, Labour would not have been defeated and so John would not be sad.

---

49. Lewis, D., *Counterfactuals*, p.33. 50. Note on p. 101.

Mackie's solution to the problem is to analyse 'X is necessary-in-the-circumstances for and causally prior to Y' as follows:

> Consider a world like the actual one up to the time of the proposed cause X, suppose X does not occur and let the world run on; if Y does not occur X was necessary-in-the-circumstances for and causally prior to Y. [51]

Making the appropriate modification to *Nec 2*[52] one has:

*Nec 3** 'X causes Y' is analysed as:

(1) X and Y both occur and are distinct existences.

(2) In circumstances <no later than X; distinct existences from X> if X does not occur then Y does not occur.

There is an objection to *Nec 3** which I shall now discuss.

Objection.[53] An amoeba called Jojim splits into two ameobae Jo and Jim. Let C be the quasi-event that Jojim splits at time t_o (into some pair of amoebae).

Let A be the quasi-event that Jo is present at some time after t_o.

Let B be the quasi-event that Jim is present at some time after t_o.

Then A and B are distinct existences and, even in the circumstances in which Jojim splits, A is necessary for B. For if Jo is not present and Jojim splits then Jojim splits into two amoebae which are not the same as Jo and Jim, so Jim is not present either. Yet A does not cause B.

Reply: This is not, I claim, a problem of collateral effects, but rather one to do with distinct existences. A modification needed to handle the case of Joanne giving birth resulting in Harry becoming a father[54] also shows that the distinct existences requirement is not satisfied in the Jojim example. The required modification is that Y

---

51. Mackie, J.L. *The Cement of the Universe,* p.51.
52. See p. 73.
53. Due to Edgar Sleinis.
54. See p. 44.

be distinct from the conjunction of X and earlier quasi-events. Using this modification, one has

Nec 3 'X causes Y' is analysed as:

(1) X and Y both occur and Y is a distinct existence from the conjunction of X and all quasi-events occurring before X, *and*

(2) In circumstances <no later than X; distinct existences from X>, if X does not occur then Y does not occur.

5.2 Collateral effects also present a difficulty for the Sufficiency Thesis. Suppose a vending machine contains chocolate bars and packets of peanuts. At random the machine sometimes goes into a short-lived state S which inevitably results some time later in the production of a packet of peanuts *without any changes in the meantime*. Unless the machine is in state S one cannot push the shilling into the slot. If the machine is in state S and a shilling is put into the slot the chocolate bar soon comes out (sooner than the free peanuts). Producing the chocolate bar would now, using the analysis *Suff 1* or *Suff 2*, be sufficient-in-the-circumstances for producing the peanuts. But the chocolate bar coming out does not cause the peanuts to come out. Thus one has a counter-example to the sufficiency of the Sufficiency Thesis, *Suff 2*. I propose instead *Suff 3:*

'X causes Y' is to be analysed as:

(1) X and Y both occur and Y is a distinct existence from the conjunction of X and all quasi-events occurring before X, *and*

(2) In circumstances <occurring no later than X; distinct existences from X>, if X occurs Y is highly likely to occur.

5.3 I shall now consider Lewis' proposed solution to the problem of collateral effects, which he calls the *problem of epiphenomena*. Lewis denies the truth of the counterfactuals which are commonly thought to raise difficulties. Suppose c causes both e and f, and suppose that e is earlier than f but e does not cause f. Lewis claims that if event e had not occurred, c would still have occurred but would have *failed* to cause e; he says "If e had been absent, it is not that c would have been absent ... Rather, c would have occurred just as it did but would have failed to cause e."[55] Lewis justifies this assertion on the grounds that c would have been determined by earlier events, so the absence of c would also involve a departure from a law of Nature. There is therefore less difference between the actual and possible world if *one* law is broken and *one* event changed than if *one* law is broken and *two* events are changed. Lewis says "I shall be content, for now, if I can give an analysis of causation that works properly under determinism"[56] so it is, perhaps, not a fair criticism to point out that his analysis fails in cases in which Determinism is not assumed. Consider the following example. A radiocative nucleus N_1 decays, spontaneously, into a nucleus N_2 (event c) emitting an α-ray which is then present in the vicinity of N_2 for a short period (event e). It is inevitable that within ten minutes N_2 decays into N_3, emitting a β-ray (event f). In this case, event e does not cause event f. On Lewis' account, in order to modify the actual world least so as to ensure that e does not occur, one states that N_1 decays into N_2, and hence into N_3, emitting the β-ray; but N_1 decays into N_2 without an α-ray being in the vicinity of the nucleus N_2. This involves breaking some law of Nature. However, one might simply modify the

55. Lewis, D., "Causation", *Journal of Philosophy*, vol.70(1973), p.566.

56. *Ibid.*, p.559.

world so that the spontaneous event of N_1's decay did not occur. Thus one preserves *one* law at the cost of changing *two* events. Now in the case of causal dependence one has the Lewis-conditional 'If ~O(c) then ~O(e); in this case it seems that the least change to the actual world is obtained by *preserving* the law and changing one event, rather than by *departing* from the law and having e still occur. Thus in this example Lewis' solution involves a rate-of-exchange in which a law is worth more than one event, but is worth less than two events. This is somewhat bizarre.

However, there is another objection to Lewis' account. Assume that there are, which is conceivable, determining laws such that the past determines the future and vice versa. Then in a straightforward case of causal dependence the Lewis-conditional 'If ~O(c) then ~O(e)' involves the claim that some world in which O(c) and O(e) are both false is more like the actual world than any world in which O(c) is false but O(e) is true. But by preserving the law in order to make e not occur, one *either* breaks the law that if e occurs some effect f of e also occurs, *or* one changes some *infinite* sequence of future events.

On the one hand it seems that breaking past instances of laws is considered to make a greater difference than breaking future instances. On the other hand, a law is worth infinitely many events; why not, then, in the case of epiphenomena, change c and the whole infinite causal chain leading up to c, rather than break the law that if e does not occur, c does not occur? In either case it is clear that changes in the past must be considered to make a greater difference to possible worlds than changes in the future. Rather than assuming this principle, it seems better to simply stipulate, as in

Mackie's solution, that the past is not to be changed. In that case we have a convergence between the proposed modifications to the accounts of Mackie and Lewis. On Lewis' account we require O(c), O(e) and the Lewis-conditional 'If events actually occurring no later than e (other than c) still occur and ~O(c), then ~O(e)'.

5.4 Thus I claim that, whether one is using a Mackie-type Necessity Thesis, a Mackie-type Sufficiency Thesis, or a Lewis-type account, the history before the causal situation needs to be considered. Hence, at least in some cases, the sentence 'X causes Y' is made true not just by what happens in the time interval between X and Y, but also by events preceding X. In other words, it is conceivable that there be two otherwise identical situations of which one is causal but the other is not, which differ only in the history before the events being considered. This conclusion is,perhaps, incompatible with commonly accepted pre-philosophical theories of causation involving, say, *causal powers*, but this incompatibility is not surprising because such theories are versions of *Causal Realism*.

Section Six: The Problems of Underdetermination and Overdetermination

6.1 The problem of *underdetermination* is that sometimes a quasi-event A causes *some part* of another quasi-event B but A is causally irrelevant to the *remainder* of B. In such cases A is *not* said to cause B.

Example One: One would not say that the fire causes the horizontally moving smoke. The fire causes the smoke but the wind makes the smoke move horizontally.

Example Two: One would not say that cyclone Alice caused the heavy

rainfall throughout Queensland for the year, if cyclone Alice were only responsible for the heavy rainfall in the north of the state and if other factors caused the heavy rainfall further south.

The problem of underdetermination is not merely a difficulty for the advocate of the sufficiency of the Necessity Thesis. If cyclone Alice were to occur in December and there had been heavy rainfall in the south of Queensland in January and February then cyclone Alice might be both necessary and sufficient in the circumstances for the heavy rainfall throughout Queensland for that year.

My solution to the problem of underdetermination is to require that X be necessary (sufficient) in the circumstances for every *regular* part of Y.[57]

Note: Someone might insist that the sentence 'The fire causes the horizontally moving smoke' is true. In that case the sentence should be paraphrased as 'The fire causes the smoke and the smoke moves horizontally'.

6.2 An analysis of causation must distinguish between those cases of *overdetermination* of C by A and B in which one says that A caused C, and those cases in which one does not. Cases of overdetermination will be classified into *simple* and *alternative* overdetermination, and alternative overdetermination will be classified into symmetric and *asymmetric* cases. I shall also distinguish between *temporal* asymmetries and *non-temporal* asymmetries. Thus there is a classification of cases of overdetermination into the following *four* kinds:

(1) Simple Overdetermination

For example, consider the causal sentence 'Pouring a litre of

57. See p.43 for the definition of a regular part.

paraffin on the bonfire caused it to burn'. This would be considered true even if half a litre would have been enough. Cases of simple overdetermination can be treated in the manner in which I treated 'Jane's eating the open jar of fishpaste caused her feeling ill', where the speaker does not wish to commit himself about the precise cause but is willing only to indicate an *area* in which the cause is to be found. Thus one could paraphrase 'Pouring a litre of paraffin on the bonfire caused it to burn' as:

> There is some positive quantity x no greater than a litre such that the quasi-event described by 'At least quantity x of paraffin was poured onto the bonfire' caused the bonfire to burn.

An alternative solution to the problem of simple overdetermination proposed by Mackie is to treat the proposed cause as a unit. Thus Mackie, in describing the case of a chestnut hit by a hammer, says "... we either plug in the hammer blow as a whole or leave it out as a whole".[58] Now one might be tempted to claim that, relative to the causal field being considered, one does not distinguish between, say, a litre and half a litre of paraffin. But this is not correct. For if one asked the question 'What caused the bonfire to burn?' one might consider the following reply to be relevant 'The pouring of a litre of paraffin on it did the trick, and it really needed a whole litre'. Perhaps one could defend Mackie's solution by claiming that some quasi-events admit of *degrees*. If a quasi-event A differs from all the norms in a qualitatively similar but quantitatively different way from quasi-event B, A and B might be considered different degrees of the same quasi-event. Such a defence, involving the notion of the degree of a quasi-event, is unnecessary and I shall

---

58. Mackie, J.L., *The Cement of the Universe*, p.44.

not investigate it in greater detail.

(2) Symmetric Alternative Overdetermination

Mackie gives the following examples:

> (i) A man is shot dead by a firing squad, at least two bullets entering his heart at once, either of which would have been immediately fatal;
>
> (ii) Lightning strikes a barn in which straw is stored, and a tramp throws a burning cigarette butt into the straw at the same place and at the same time; the straw catches fire.[59]

In such cases there is no way of deciding which quasi-event should count as the cause, so perhaps one should call neither the cause. (Although as I said in Section Three, one is also reluctant to deny that either is the cause.) If one does adopt the Necessity Thesis one would say that in such cases the genuine cause is the quasi-event obtained by taking the disjunction of the two proposed causes.

(3) Temporal Asymmetries

Consider the following example. Someone with tuberculosis takes a drug and goes to live in a hot dry climate. Either would be sufficient to cure him. The drug cures him in two weeks but the change of climate would have taken six months to cure him. In this case, even though taking the drug is not necessary in the circumstances for the cure, it is nonetheless the *cause* of the cure.

Such cases can be handled using the method Mackie proposes for all cases of *asymmetric* overdetermination. In the above example one can claim that the drug is necessary for a *cure within a month*. As in the case of 'Jane's eating fishpaste in the pantry caused her

59. Mackie, J.L., *The Cement of the Universe*, p.44.

feeling ill', there are various possible paraphrases of 'Taking the drug cured the man';

Paraphrase One: Taking the drug caused the quasi-event described by 'The man recovered health'.

Paraphrase Two: Taking the drug caused the quasi-event described by 'The man recovered health in two weeks'.

Paraphrase Three: The disjunction of Paraphrases One and Two.

Using the analysis *Nec 3*,[60] Paraphrases Two and Three would be true. I suggest that Paraphrase Three is the most accurate; one does not state whether the time taken for the man to recover health is part of the description of the effect or not.

6.3 (4) Non-temporal Asymmetries

There are various examples of *non-temporal asymmetries* in which both A and B occur and either by itself would ordinarily cause C to occur at the same time; one says A causes C, but B does not. I first discuss the difficulties these examples raise for the necessity of the Necessity Thesis.[61]

Example One:

> ... conditions (perhaps unusual excitement) plus constitutional inadequacies are present at 4.0 p.m. that guarantee a stroke at 4.55 p.m. and consequent death at 5.0 p.m.; but an entirely unrelated heart attack at 4.50 p.m. is still correctly called the cause of death, which, as it happens, does occur at 5.0 p.m.[62]

In this example I assume that the "guarantee" is not absolute, and that there is no stroke - perhaps the heart-attack *prevents* the

---

60. See p.88.

61. I ignore the problem of overdetermination in the case of the Sufficiency Thesis, because the context in which the problem arises is that in which the question 'What caused Y?' is asked.

62. Mackie, J.L., *The Cement of the Universe*, p.44.

stroke. Another example is a slight modification to Mackie's example (iv).

Example Two:

> Smith is sent to shoot the president as he comes out of a certain building at 2.00 p.m. Smith is watched by Brown who, if he failed to see Smith in place at 1.50 p.m., would himself shoot the president at 2.00 p.m. However, Smith is in place and shoots the president.

In this case the president would have been shot anyway (at the same time). For one is to assume from a knowledge of Smith's character, that if he proceeds according to plan until 1.50 p.m. he will shoot the president, but if Smith does not proceed according to plan until 1.50 p.m., Brown still has time to shoot the president at 2.00 p.m. In this second example, one could replace Smith by a mechanical device such as a time-bomb. Brown, then, checks that the time-bomb has not been discovered. Although in this example it might be essential that Brown be considered an agent, the cause that actually occurs is a transeunt cause.

I suggest that all cases of non-temporal asymmetries have in common the breaking at some point of one of the two potential causal chains. The stroke does not occur, and the failure of Smith to be in place does not occur. Thus in both cases a rival causal chain is broken.

One might think that *temporal asymmetries* are also of this kind. For example, a detailed description of the potential cure by a change in climate might, perhaps, presuppose that the person still suffers from tuberculosis after one month. In actual cases this may be so, although it is debatable. However, assuming that mnemic causation is conceivable, there are *hypothetical* examples in which the temporal

asymmetry involves no break in a causal chain, because there is no causal chain. For example, suppose a magic ritual would, without any intermediate changes, suddenly cure the person six months later. (The continued sickness of the person is not here presupposed because should someone be cured and yet again contract tuberculosis, the ritual still cures him.) In that case one would nonetheless say that the drug, not the ritual, cured the sick person.

Mackie suggests that the solution to the problem of non-temporal asymmetries is that the cause was "necessary in the circumstances for the result as it came about."[63] While I think Mackie is basically right, I reject his clarification of this solution in terms of the distinction between facts and events as effects. It seems Mackie would claim that the result of the heart-attack is the *fact* that the person died of heart-attack and that because the *event* of the person dying of heart-attack is the same *event* as the person dying, one has a case of event-causation in which the heart-attack causes the death. My grounds for claiming that this is Mackie's position are that in a more complicated example ((v), discussed below) Mackie describes the result as the fact that the traveller died of thirst. It is clear that in the present example the person did not die of a stroke - he did not have one. But one should not describe the death as *death by heart-attack* until one has decided that the heart-attack caused the death. However, there seem to be two suitable descriptions. Perhaps - as I think Mackie intends - one is to describe the death as death as a result of the *post-heart-attack condition* (which occurs at 4.55 p.m.), thus putting into the description undisputed causes later in the chain. Alternatively, one might describe the death as *death which is not the result of a stroke*. The former description does

63. Mackie, J.L., *The Cement of the Universe*, p.46.

not cover the conceivable case in which the heart-attack mnemically causes death ten minutes later without any intermediary change. However, the use of the latter description presents a new problem in the case of symmetric overdetermination. In Example (ii)[64] the lightning-stroke is necessary in the circumstances for the straw to burn *not as a result of the throwing down of a cigarette butt.* For even without the lightning stroke the straw would have burnt as a result of the throwing down of the cigarette butt. Thus in the description of the effect it would have to be stipulated that the alternative cause does not occur, rather than that the alternative causal *relation* does not occur. So, in Example One above,[65] the relevant description of the death would have to be 'death without a stroke occurring earlier' rather than 'death not caused by a stroke'. The correct paraphrase of 'The heart-attack caused the death' would then be 'The heart-attack caused the quasi-event described by 'death occurred and no stroke occurred earlier' '. But in this case one has an example of *underdetermination* so the causal sentence would be false. Thus there are difficulties with Mackie's solution.

6.4 Lewis proposed to solve the problems of asymmetric overdetermination by considering causal chains. Suppose c_1 causes e and that c_2 also occurs and does not cause e but would have caused e if c_1 had been absent. Say c_1 is the event of the heart attack, c_2 the pre-stroke conditions, and e the death. Lewis then suggests that there is an intermediate event d such that e depends on d and d depends in turn on c_1.[66] But I am assuming the coherence of mnemic causation, so it is conceivable that the heart attack directly causes the death

---

64. See p. 94.
65. See p. 95.
66. Lewis, D., "Causation", *Journal of Philosophy*, vol.70(1973), p.567.

without intermediary changes. However, a variant of Lewis' solution does succeed.

A first attempt at a suitable analysis might be to require that the circumstances include all quasi-events occurring earlier than Y, or no later than X. But it then follows that if X causes Z which causes Y, X is no longer necessary in the circumstances for Y. So I propose instead the following account:

Nec 4 Y is said to be *causally dependent on X* if

(1) Y is a distinct existence from the conjunction of X and all quasi-events occuring before X, *and*,

(2) If Z is any regular part of Y then in circumstances <occurring no later than X or occurring earlier than Y; distinct existences from X and Y> if X does not occur Z does not occur.

Then 'X causes Y' is analysed as

> There is a finite chain of quasi-events $X=X_1, X_2, \ldots, X_n = Y$ such that X_{r+1} is causally dependent on X_r.

I shall not provide a causal chain account for the Sufficiency Thesis since it is possible that the probability of the occurrence of X_{r+1} given the occurrence of X_r, $r = 1 \ldots (n-1)$, is high, but the probability of X_n given the occurrence of X_1 is low.

Using *Nec 4*, if the heart attack is followed by death without intermediary changes death is causally dependent on the heart attack, for one is supposing that the stroke does not occur. Similarly, in the example of Smith shooting the president, in the circumstances in which Brown does not shoot the president it is necessary that Smith shoots the president. Finally, consider again Mackie's example (v):

> A man sets out on a trip across the desert,
> He has two enemies. One of them puts a

> deadly poison in his reserve can of drinking water. The other (not knowing this) makes a hole in the bottom of the can. The poisoned water all leaks out before the traveller needs to resort to this can; the traveller dies of thirst. [67]

In this example there is some doubt about the cause of death. One might judge, like Mackie, that the cause is the puncturing of the can, or one might judge that the poisoning of the water and the puncturing of the can overdetermine the traveller's death. It is instructive to see the way in which the analysis *Nec 4*, reflects this doubt. *On the one hand*, one might divide the facts into the following quasi-events - or absence of quasi-events:

A: The poison is put into the can
B: The can is punctured
$\overline{C}$: No poison is drunk
D: There is no water in the can
E: The man dies

One then argues that E depends causally on D, for the circumstances include $\overline{C}$. D depends causally on B. So B causes E. *On the other hand*, one might divide the facts into the following quasi-events:

A: The poison is put into the can
B: The can is punctured
$\overline{C}$: No poison is drunk
D*: There is no proper drinking water
E: The man dies

In this case, E depends causally on D* which is overdetermined by A and B. So A and B overdetermine E.

Notice that D and D* are not distinct existences so one cannot combine the two ways of looking at the situation. Notice also that even if $\overline{C}$ is not a quasi-event but it is part of every normal state-of-affairs it is still part of the circumstances.

---

67. Mackie, J.L., *The Cement of the Universe*, p.44.

Note: The account is easily modified to exclude *mnemic* causation and *action-at-a-distance* if one so wishes. One simply stipulates that if Y depends causally on X, X and Y must be in spatio-temporally adjacent or overlapping regions.

Section Seven: Simultaneous Causation

7.1 It is perhaps debatable whether simultaneous causation is conceivable. If it is not, then it is easy to stipulate in an analysis of 'X causes Y' that X is earlier than Y. So, in order to make my account adequate in both cases, I shall assume that simultaneous causation is *conceivable*. The analyses *Nec 4* ans *Suff 3* permit simultaneous causation. *Suff 3* is obviously inadequate as a sufficient condition if X and Y are simultaneous, for Y would be part of the circumstances and so any two simultaneous events would cause each other. Hence *Suff 3* is modified to obtain *Suff 4:*

'X causes Y' is analysed as:

(1) X and Y occur and Y is a distinct existence from the conjunction of X and all quasi-events occurring earlier than X

and

(2) In circumstances <no later than X; distinct existences from X and Y>, if X occurs Y is highly likely to occur.

The difficulty with both *Nec 4* and *Suff 4* is that they allow cases of mutual simultaneous causation. In this section I provide an extra necessary condition for X to cause Y, in the case of simultaneous causation. I discuss *Nec 4*, but this extra condition could also be used with *Suff 4*.

Consider the following examples:

Example One:

> If I view as a cause a ball which impresses a hollow as it lies on a stuffed cushion, the cause is simultaneous with the effect. [68]

Example Two:

> From the fact that a bar of iron is now glowing we can certainly infer ... that it is now at a temperature of 1000° or over. Yet we should not say that its high temperature was caused by the glowing: we say that the high temperature causes the glowing, not vice versa. [69]

These examples seem to be actual causes of simultaneous causation.

Now someone might object that in fact there are no cases of simultaneous causation, but that if one told a sufficiently detailed story about, say, the motion of molecules it would be seen that the cause always precedes its effect. However, it is incontestable that the above are examples of *causation*. What is disputed is whether in these examples cause and effect are *simultaneous*. Now even if the cause and the effect are not simultaneous, this lack of simultaneity is *not directly observed* but is *inferred* from some theory about microscopic or sub-microscopic events. This theory was discovered *a posteriori* so it is conceivable that the theory is *false*, and so it is also conceivable that cause and effect are simultaneous. Thus, even if the examples are not *actual* cases of simultaneous causation, they can be used to show that there are coherent *hypothetical* cases of simultaneous causation.

Suppose Y depends causally on X and X and Y are simultaneous; in that case, using the analysis *Nec 4* one has:

---

68. Kant, I., *Critique of Pure Reason*, B248.
69. Gasking, D., "Causation and Recipes", *Mind*, vol.64(Oct.1955) p.480.

> In circumstances <occurring no later than X; distinct existences from X and Y> if X does not occur Y does not occur.

Suppose not only that in circumstances <occurring no later than X; distinct existences from X and Y> if X does not occur Y does not occur but that also in these circumstances if Y does not occur X does not occur. Then, according to *Nec 4*, X and Y cause each other, but causation is generally considered to be an *anti-symmetric* relation.

Now in the two examples of simultaneous causation given above one can argue that this symmetry does not arise. Consider the example of the ball and the depression in the cushion. The circumstances include the past motion of the ball and the past state of the cushion. Presumably the ball is about to drop onto the cushion and, presumably, the cushion has no depression. Then if there is no ball on the cushion, presumably there is no depression. But if there is no depression one of two anomalies might have occurred: the ball might have stopped, suspended just above the cushion; or the ball might be on the cushion without causing a depression. The inability to decide which anomaly has occurred prevents one correctly asserting the conditional 'In the circumstances, if there had been no depression there would be no ball on the cushion'. At any rate, if one gives a *meta-inference* account or a *rationality-judgment* account of conditionals, the choice between two anomalies prevents the conditional being true.

Again, consider Example Two (the iron glowing at 1000°). One is to assume as part of the circumstances that a process has been going on, which is sufficient in the (earlier) circumstances for the iron to reach 1000°. There will be several such processes: these

could be called 'heating processes', but that is to use a causally-tied description. So I shall simply refer to them as H-processes where an H-process is by definition one of processes $P_1 \ldots P_k$, where it happens that $P_1 \ldots P_k$ are all the processes which heat iron. Now, if *on the one hand* the iron is not hot and there is no glow, a departure[70] from *one* non-accidental generalisation has occurred (namely, that the H-process makes the iron hot), but if the iron is not hot and there is a glow there have been *two* departures from non-accidental generalisations (for it is also a non-accidental generalisation that hot iron glows). Hence, in the circumstances, if the iron is not hot it does not glow. *On the other hand,* in the circumstances, if the iron does not glow then *either* there is a departure from the non-accidental generalisation that hot iron glows *or* there is a departure from the non-accidental generalisation that the H-process makes the iron hot. So, as in the Example One, the choice between two anomalies prevents the correct assertion of the conditional 'If the iron had not glowed, it would not have been hot'.

7.2 There are, however, several objections to *Nec 4* based on other examples of simultaneous causation.

Objection One:[71] In the above two examples it was assumed that the proposed cause itself had a determining cause. But suppose that on 50% of occasions balls stopped moving, and suppose that on 50% of occasions the H-processes fail to make the iron hot. Then one *would have* grounds for choosing between one anomaly rather than the other. So in the circumstances if the depression does not occur,

---

70. A non-accidental generalisation is not necessarily universal, hence by a departure from it I mean an exception to the corresponding universal generalisation.

71. This objection is due to Michael Tooley.

the ball is not on the cushion. Again, in the circumstances, if the iron does not glow, then it is not hot.

Reply: In such cases the analysis *Nec 4* fails. It is therefore replaced by a more complicated account of causal dependence which is based on the requirement that if in fact Y does not cause X and if one had removed some cause of X *earlier* than X, but one had independently ensured that Y occurred, then X would not have occurred. In other words, some quasi-event is necessary in the circumstances for X, even if the occurrence of Y is part of the circumstances. My justification for this kind of account is that the aim of my conditional analysis is to state in what respects cases of causation are like their *manipulability paradigms*. In the case of the ball and the cushion, one could not ensure that there is a ball by merely producing a depression in the cushion. To avoid description in terms of agents one simply stipulates that the circumstances be adjusted by removing some cause of X. Here I assume that *Nec 4* gives an adequate account of causation if X is *earlier* than Y, consequently there is no circularity in mentioning the cases of causation in which the cause is earlier than its effect in the analysis of *simultaneous* causation. Accordingly one obtains the analysis *Nec 5* by making the following proviso:

Suppose that if X and Y are simultaneous and, using *Nec 4*, it would seem that X and Y depend causally on each other, then an extra condition (EC) is required if Y is to be *causally dependent* on X, namely:

(1) There is at least one quasi-event Z earlier than X on which (using *Nec 4*) X depends causally.

and

(2) If Z is *any* quasi-event earlier than X on which (using *Nec 4*) X depends causally, *then*:

in circumstances <occurring no later than X; distinct existences from X, Y and Z>

if Z does not occur but Y occurs then X does not occur.

In the modified *hot iron glowing* example one considers circumstances in which the iron is not subject to the H-process. I suggest that the relevant non-accidental generalisations are:

(1) Some H-process has occurred whenever iron is hot;

(2) If iron is hot it glows;

(3) If iron glows then it is hot;

but *not*

(4) Some H-process has occurred whenever iron glows.

If one supposes that the iron is hot and glows, then there are departures from (1) and (4). If one supposes that the iron is hot but does not glow, then there are departures from (1) and (2). Since (4) is not a non-accidental generalisation, in the special circumstances where there has been no H-process the iron being hot is sufficient for it to glow.

If no H-process has been used and one supposes the iron glows yet is not hot, then there are departures from (3) and (4). But if the iron glows and is hot there are departures from (1) and (4). In this case one cannot choose between the two anomalies, so one should not say that in the special circumstances the glow is *sufficient* for the iron to be hot. Hence by *Nec 5* the heating of the iron causes the glowing.

Note One: It might seem peculiar that (1) and (3) but not (4) are non-accidental generalisations where (1) and (3) jointly *entail* (4). This peculiarity will be discussed in greater detail subsequently; here I merely point out that both the modal operators L_1 and L_2 defined by:

L_1 (All P are Q) $=_{Df}$ (All but perhaps one P is a Q)

L_2 (All P are Q) $=_{Df}$ (It is highly probable that a P is a Q)

have the property that if p and q jointly entail r, then Lp and Lq need not jointly entail Lr.

Note Two: It suffices to argue that the generalisation (4) has a *lower degree of fundamentality* than the generalisation (1). (Generalisation (4) could still be deemed non-accidental.)[72]

Note Three: Because I am providing here an account of causation in terms of the notion of *non-accidental generalisations*, it is interesting - though not crucially important - to discover why one does not say that the generalisation 'An H-process has occurred whenever iron glows' is a law of Nature, or, more precisely, a non-accidental generalisation. The explanation, I suppose, for treating the generalisation 'An H-process has occurred whenever iron is hot' as a non-accidental generalistion is that it is an instance of the general principle that an H-process has occurred whenever an object is hot. But the generalisation 'An H-process has occurred whenever iron glows' is not an instance of a more general principle. This remark is of course based on Gasking's discussion, who says:

> When we have a general manipulative technique which results in a certain sort of event A, we speak of producing A by this technique ... When in certain cases application of the general technique for producing A results in B we speak of producing B by producing A ... And in such a case we speak of A causing B, but not vice versa.[73]

Note Four: My own intuitions about causation exclude cases of *mutual causation*, so in the perfectly symmetric case where neither *Nec 4* nor *EC* enables one to distinguish cause from effect, I say that neither X causes Y nor Y causes X.

---

72. See p.114 for an account of degrees of fundamentality.

73. Gasking, D., "Causation and Recipes", *Mind*, vol.64 (Oct.1955), p. 483.

Note Five: To obtain a version of the Sufficiency Thesis which handles cases of simultaneous causation, one could add the extra condition *EC* to *Suff 4*, thus obtaining *Suff 5*.

Objection Two: In perfectly symmetric cases one would say 'Either X caused Y or Y caused X; I do not know which', rather than 'Neither X nor Y caused the other'. For example, in a world where iron being hot and iron glowing always occurred together but the conjunctive quasi-event that the iron is hot and glowing had no cause, one would say that either the heat caused the glow or the glow caused the heat.

Reply: One might be *tempted* to say that if events of kinds *C* and *E* are correlated but are otherwise spontaneous, then the *C* cause the *E* or vice versa; but this, I suggest, is not a matter of English usage so much as a piece of pre-philosophical Metaphysics. Suppose there is some *actual* situation where spontaneous events are correlated; say it is discovered that whenever a radium atom decays then some nearby radioactive lead atom also decays and vice versa. If, in that case, a *reliable* person claimed that he knew which was the cause and which was the effect, one might entertain the hypothesis that there *is* an irreducible element of causal priority "in the objects" which this person intuits. But the claim that one causes the other but no one knows which obviously seems to be based on a belief in *Causal Realism*.

Objection Three: The account given of causation is now extremely complex. Various ad hoc rules are introduced in order to meet objections. Consequently the account is not plausible.

Reply: My aim is to give an account of those respects in which all cases of causation are like the paradigms based on manipulability. That the account is complex is to be expected; a similar account for

tables would also be complex. A sound objection to the account would be made if the features were *so* complex that it was *inconceivable* that a human being could base the judgment of similarity with the paradigms on these features. This degree of complexity has surely not been reached. Furthermore, the account does not involve *ad hoc* rules. Rather, in most cases it suffices to consider the features of causal situations such that one can prevent the effect by preventing the cause or produce the effect by producing the cause. But in some cases of simultaneous causation in which X causes Y, one has to consider the features of the causal situation such that if one takes all steps possible before the cause X in order to prevent X but otherwise ensures that Y occurs, then the cause does not occur. The accounts *Nec 5* and *Suff 5* are based on these features.

PART TWO

Section Eight: Non-accidental Generalisations

8.1 In Section Nine I shall argue that the conditionals required for the analysis of causation are *meta-inference* conditionals based on the correct non-accidental generalisations. In this section I compare non-accidental generalisations with *laws of Nature*. Laws of Nature are usually contrasted with accidental universal statements, the difference being that in some sense laws are necessary universal statements. For instance, Nagel says:

> No one seriously disputes the claim that a distinction something like the one baptised by the labels 'accidental' and 'nomic' universality is recognized in common speech and in practical actions. The question in dispute is whether the prima facie differences ... require the acceptance of the "necessity" associated with universals of law as something "ultimate" or whether nomic universality can be explicated in terms of notions that are less opaque. [74]

Thus a law of Nature is of the schematic form Lp where L is the modal operator 'of nomic necessity' and p is a universal generalisation such as 'All P are Q'.

For two reasons I want to avoid considering laws of Nature in my account of the conditionals used in causation.

Reason One: I am not convinced that all the generalisations which license the counterfactuals used in the analysis of causation are nomic *universal* generalisations (laws of Nature). For, I claim, some non-physical entity, such as the human mind or soul, might directly cause some physical event - this would be a case of immanent causation. Accordingly, suppose that Libertarianism is correct, and suppose also

74. Nagel, E., *The Structure of Science*, p.53.

that when a free choice is made something happens in the brain which is an *exception* to those generalisations that are usually considered to be laws. Then the "laws" are not *universal* generalisations, but nonetheless they license counterfactuals and are consequently not accidental. Nor can the universal nature of the "laws" be ensured by making them *probabilistic*. It is possible that the only exceptions to the "laws" occur when someone makes a free choice in the libertarian sense of a free choice and so, apart from exceptions due to immanent causation, the non-probabilistic "laws" are universal.

Reason Two: One sometimes supposes a hypothetical situation in which there are exceptions to what are, perhaps, universal generalisations. For example, suppose that a thousand years ago a stone, for no reason, remained ten feet above the earth for two minutes and then fell. Suppose also that (now) a stone is thrown upwards. Surely, within the scope of these suppositions, one should assert that the stone falls to the earth. Yet one has supposed that the "law" that heavy unsupported objects fall to the earth has been broken and so the "law" is not a nomic *universal* generalisation. From what, then, does one infer that the stone falls to the earth? There are two answers someone might give.

Answer One: One uses induction. On most occasions in the past unsupported heavy objects fell to the earth, so on this occasion also the stone falls to the earth.

Answer Two: One still assumes that it is a *non-accidental but not necessarily universal generalisation* that unsupported heavy objects fall to the earth, and from that assumption one infers that the stone falls to the earth.

It is not within the scope of my thesis to provide an account

of induction. (By 'induction' I do not mean any non-deductive inference but simply ordinary induction, of which the paradigm is 'All the crows I have seen have been black so the next one I see will be black'.) I shall merely claim, without argument, that to use induction commits one to the belief that the regularity being projected is not accidental and so can be treated as a non-accidental generalisation. But is this generalisation a probabilistic law? I suggest it is not. If asked what would happen if, a thousand years ago, a stone remained ten feet above the earth, one does not suppose a new universe with new laws. Rather, one uses the actual "laws" to infer from the supposed circumstances that the stone which is now thrown upwards, falls back to the earth. But one cannot make reliable inferences from self-contradictory premises. The solution to this difficulty is to accept Answer Two, namely, that the law is converted into a non-accidental generalisation from which it is inferred non-deductively that the stone which is now thrown upwards will fall to the earth. In particular, I consider that Nagel's meta-inference account of counterfactuals needs to be modified. He says:

> A counterfactual can be interpreted as an implicit metalinguistic statement ... asserting that the indicative form of its consequent clause follows logically from the indicative form of its antecedent clause, when the latter is conjoined with some law and the requisite initial conditions for the law.[75]

However, one would accept as correct the counterfactual 'If a thousand years ago a stone had, for no reason, stayed two minutes above the earth, and if (now) a stone were thrown upwards then this stone would fall to earth', in which the conjunction of the antecedent clause with the law that all heavy unsupported objects fall to the earth is self-

---

75. Nagel, E., *op.cit.*, p.72.

contradictory.

8.2 The Logic of Non-accidental Generalisations

Typically a law, that is a nomic universal generalisation, is of the form $L((x_1)(x_2)\ldots(x_n)F(x_1,x_2,\ldots x_n))$ that is 'Necessarily for all x_1 for all x_2, ... for all x_n, $F(x_1,\ldots,x_n)$' where $F(x_1,\ldots x_n)$ is a formula with free variables $x_1\ldots x_n$, no bound variables, and no proper names, and L is the modal operator 'of nomic necessity'. One could write $L((x_1)\ldots(x_n)F(x_1,\ldots,x_n))$ as $LUF(x_1,\ldots,x_n)$ where LU is the combination of modal operator and as many universal quantifiers as are required.

For a non-accidental generalisation, instead of the operator LU one has the operator NA where $NAF(x_1,\ldots,x_n)$ is an abbreviation for 'It is not an accident that for almost all $x_1,\ldots,$ and for almost all $x_n, F(x_1,\ldots,x_n)$'. If one supposes that a sentence p is true and p does not entail any "departure" from the non-accidental generalisations, one can use the rule of inference:

$$NAF(x_1, \ldots, x_n)$$
$$\therefore\ F(a_1, \ldots, a_n);$$

where $a_1,\ldots,a_n$ are either names of individuals or unique referring expressions.

If one can then deduce q from the conjunction of p and $F(a_1,\ldots,a_n)$, one can assert the meta-inference conditional 'If p then q'. In this situation there is no departure from Nagel's account, in which one deduces q from p and the laws of Nature.

However, when p does entail "departures" from the non-accidental generalisations, another form of inference is required. One might attempt to use Probability Theory to provide the inferences. But, as I have already argued, $NAF(x_1,\ldots,x_n)$ is not to be considered a

probabilistic statement. For example, suppose F(x,y) is:

> If x and y are two objects they attract each other with a force proportional to the product of their masses and inversely proportional to the square of their distance apart;

NAF(x,y) *is not to be interpreted as:*

> If x and y are two objects it is highly likely that they attract each other with a force proportional to the product of their masses and inversely proportional to the square of their distance apart.

However, one might want to rank non-accidental generalisations - some are *more fundamental than others*. So one might assign some numerical strength to NAF($x_1,\ldots,x_n$) and thus obtain 'Deg(NAF($x_1,\ldots,x_n$)) = z' as an abbreviation for:

> The non-accidental generalisation NAF($x_1, \ldots, x_n$) has a degree z of *fundamentality*.

In the example discussed in reply to Objection One in Section Seven,[76] I considered four generalisations, corresponding to the conditionals:

(1) If a piece of iron is hot, it has been subjected to an H-process.

(2) If a piece of iron is hot, it glows.

(3) If a piece of iron glows, it is hot.

(4) If a piece of iron glows, it has been subjected to an H-process.

I suggested that (4) was not a non-accidental generalisation. But perhaps it would be more correct to say that (4) is of a lower degree of fundamentality than (1), so when exceptions to (1) and (4) are contrasted with exceptions to (1) and (2) one has merely to compare

---

76. See pp. 104-106.

the degrees of fundamentality of (2) and (4). Now (4) is derived from (1) and (3). So, if (2) and (3) have the same degree of fundamentality, one might expect (4) to have a lower degree of fundamentality than (2). Thus one obtains the counterfactual 'If there had been no H-process but the iron glowed then it would not be hot'. In this way, by considering the degrees of fundamentality of the non-accidental generalisations, one can formulate a method of non-deductive inference from a sentence p and the non-accidental generalisations even when p entails some "departure" from the non-accidental generalisations. In situations where no methods of non-deductive inference are permitted from p and the non-accidental generalisations other than that indicated in this section, I call the meta-inference conditional a *restricted* meta-inference conditional. Restricted meta-inference conditionals provide, I suggest, the appropriate modification of Nagel's account for situations in which one is to suppose "exceptions" to the laws of Nature.

Note: I do not propose to give an analysis of the notion of *fundamentality*. There is however a *mark* of the degree of fundamentality of a non-accidental generalisation, namely the *support* of that generalisation.[77] A non-accidental generalisation which *either* has very many instances *or* is derivable from another non-accidental generalisation with very many instances is said to have a large *support* and may *usually* be assumed to have a high degree of fundamentality. If a non-accidental generalisation (Nag_1) is derived from several non-accidental generalisations $Nag_2,\ldots,Nag_n$, with many instances, Nag_1 derives some support from $Nag_2,\ldots,Nag_n$ but not as much as if it were derivable from Nag_2, say, alone.

---

77. The notion of support is discussed in greater detail on pp. 166-168.

Section Nine: The Conditionals Used in the Conditional Analysis of Causation

In Section Eight I sketched an account of meta-inference conditionals based on non-deductive inferences based on the degrees of fundamentality of non-accidental generalisations. In this section I argue that such a meta-inference account of conditionals is appropriate for the conditionals used in the analysis of causation.

It has been proposed by, for instance, Mackie that one should not consider causal statements to be true or false, but as they are usually considered to be true or false it seems best to avoid giving an account of the conditionals used in the analysis of causation in terms of condensed argument or supposition/assertion conditionals. (If one were to deny that causal sentences had truth values, one would interpret my account of causation as one which gives necessary and sufficient conditions for a causal sentence to be *commonly* said to be true.) Hence I consider only *rationality-judgment* conditionals and *meta-inference* conditionals. The *apparent* advantage of the former is that one might hope to give an account involving laws of Nature rather than an account involving non-accidental generalisations. For antecedent beliefs are modified by new evidence; a (hypothetical) rational man who is assumed to have, as antecedent beliefs, the correct beliefs about laws of Nature, would modify these (correct) beliefs if he were asked to suppose, as new evidence, some situation inconsistent with the laws of Nature.

I now argue that the conditionals required for the analysis of causation are the *restricted* meta-inference conditionals discussed in Section Eight. I assume that the conditionals of the form 'If p then q' used in the analysis of causation are *rationality-judgment*

conditionals and I then show that *qualifications* have to be made which eventually render the rationality-judgment conditional indistinguishable from a restricted meta-inference conditional. Thus I assume that person P with antecedent beliefs B is given new evidence p and then judges that it is rational to believe q. I show that the following qualifications must be made.

(1) P's antecedent beliefs contain beliefs about the non-accidental generalisations.

(2) P's antecedent beliefs about the non-accidental generalisations are not necessarily those which P would rationally believe on all the past evidence. Rather, P's antecedent beliefs must contain all the correct non-accidental generalisations (that is, those which actually occur).

(3) Suppose there were new evidence p such that P should alter his (correct) beliefs about the non-accidental generalisations. In order that the analysis of causation in terms of rationality-judgment conditionals should succeed, it would *then* have to be stipulated that P's antecedent beliefs about the non-accidental generalisations are *not* altered.

(4) When P makes his judgment of rationality there is no room for any kind of non-deductive inference other than those used in *restricted* meta-inference conditionals. Furthermore, if there were room for other kinds of non-deductive inference, it would have to be stipulated that these other kinds of non-deductive inference are not used when P makes his judgment of rationality.

Step One: Suppose a powerful electro-magnet (which needs to be switched on for two minutes before it operates) is switched on and two minutes later a nearby compass needle moves. A second before the compass needle moved there was a slight earth tremor. Someone

ignorant of the laws of electricity and magnetism might rationally deny that switching the electro-magnet on caused the compass needle to move. This example shows that for the conditionals used in the analysis of causation to be rationality-judgment conditionals, P must not be ignorant of any relevant non-accidental generalisations.

Step Two: Suppose that P's antecedent beliefs were based on many earlier observations. In that case it is conceivable that there could be *accidental* universal generalisations which P rationally but incorrectly believes to be *non-accidental*. For example, suppose that the law of gravitational attraction is a *probabilistic* law: the force between two massive objects fluctuates *at random* corresponding to a small but detectable variation in the universal gravitational "constant", G. Suppose, however, that there is excellent evidence that no such fluctuations occur. Nonetheless, given a probabilistic law for the variation of G it is possible, though extremely unlikely, that by chance, in the whole (finite) universe G has been constant at all times, but at each moment it was highly unlikely that it would remain constant. Hence it is conceivable that G fluctuates, yet it is rational to believe that G is constant. Now if the conditionals used in the analysis of causation are licensed by what one rationally *believed* were laws or non-accidental generalisations, then the analysans and analysandum might have different truth-values. For example, if it is a law that G fluctuates at random, but by accident G has been constant, then it would be false that firing the rocket precisely in direction D precisely with velocity v caused it to land on the Moon. For the chance of the rocket landing on the Moon would be much the same even if the direction varied somewhat from D, or the velocity from v; it is purely an accident that the rocket

landed on the Moon. But the proposed analysandum 'Assuming the laws of Nature are as they are rationally believed to be, if the rocket had not been fired precisely in direction D precisely with velocity v, the rocket would not have landed on the Moon' would be true. Hence the antecedent beliefs B should contain the non-accidental generalisations which actually occur, not merely those in which it would be rational to believe on earlier evidence.

Again, one might argue that it is conceivable that there is a law such that if either a physical event of type P_1 or a physical event of type P_2 occurs, then there would be a pain. But, by chance, physical events of type P_2 have only occurred once and will never recur. In that case one might argue that the unique event of type P_2 caused a pain. But it would not be rational to believe in such a law unless it could not be derived from other laws. However, I have doubts about the coherence of such examples based on unique causation and therefore I rely on the first argument in Step Two to show that the conditionals used in the analysis of causation are based on the *actual* laws or non-accidental generalisations and not merely on generalisations which it is *rational* to believe are not accidental.

Step Three: Assume that, as in my version of the Conditional Analysis, the conditionals used in the analysis of causation are of the form, 'In the circumstances, if p then q'. It follows that P is asked to suppose as new evidence not only p but also the *circumstances*. In the account I have given (*Nec 5* and *Suff 5*), the circumstances include all past quasi-events, so it is unavoidable that the new evidence, which P is asked to suppose, contains all that happens before the cause. So the circumstances might contain evidence which would make it rational for P to modify his (correct) beliefs about the laws of

Nature (or non-accidental generalisations). For example, if the universal gravitational constant fluctuates at random, the past history might contain evidence which seems overwhelming that G does not fluctuate, and P would accordingly modify his (correct) antecedent belief that G fluctuates at random. Thus one would have to stipluate that P's (correct) beliefs about non-accidental generalisations are *not* altered by new evidence. But one of the characteristic differences between rationality-judgment conditionals and meta-inference conditionals (noted in Chapter One, Section Three[78]) is that in rationality-judgment conditionals the antecedent beliefs *can be modified* by new evidence.

Note: If one could give some account of when a factor is *prima facie* causally relevant, one could *restrict* the circumstances to prima facie causally relevant factors. Even with this modification there could nonetheless be circumstances containing evidence which should alter the (correct) antecedent beliefs about the non-accidental generalisations. For example, suppose that there are a hundred isolated valleys in Asia where a certain species of bamboo grows. Every twenty years all the bamboo plants flower and die. The bamboos flower first in a valley in Sikkim, they flower last in a valley in Manchuria. The valleys have been isolated for a million years. Suppose that there is an *emergent law* that when one bamboo plant flowers all others of that species flower within three months. Suppose also that there are laws that the flowering time of the bamboo plants in Sikkim varies at random from ten to thirty years, and that the bamboo plants in Sikkim are always the first to flower. So it is merely an *accident* that the bamboo plants flower every twenty years. In this

78. See Note Two on p.50.

case the flowering of the bamboo plants in Sikkim *causes* the flowering of the bamboo plants in Manchuria. But one would rationally assume that the coincidence of the flowering time is due to all the bamboo plants having a single ancestor a million years ago; that is, there is a *collateral effects* situation. Therefore one would consider the flowering times of all the ancestors of the bamboo plants in the two valleys for the past million years to be *prima facie causally relevant* and hence part of the circumstances. It follows that, even if P only supposes the prima facie relevant circumstances, P has sufficient evidence to modify his (correct) antecedent belief in the emergent law. Thus P would deny the (correct) conditional 'In the circumstances, if the bamboos in Sikkim had not flowered the bamboos in Manchuria would not have flowered [in the same year]'.

Step Four: I now argue that there is no room in the rationality-judgment conditional for any inference other than deduction and the non-deductive inferences used in the restricted meta-inference conditionals. Suppose one tried to use ordinary induction. For example, it has happened that whenever a certain species of insect has bitten a certain species of vine, the vine ceases to twine anti-clockwise and thereafter twines clockwise. If P has evidence that there have been such changes in the direction of twining and if there was room for induction, P would conclude that the biting of the insect causes the change in direction of the vine. Yet it might be a mere accident that the biting of insects is correlated with changes in the direction of the vine. So if there were room for inductive inferences, one would have to rule out the use of induction anyway, and thus one would have an argument for considering only restricted meta-inference conditionals. However, there is no room for such a

use of induction, for P believes all the non-accidental generalisations and so he is able to infer that the biting of the insect is causally irrelevant to the twining of the vine.

Suppose next that *psycho-physical parallelism* is correct; that there are *psychic* laws governing the relation between mental events, and *physical* laws governing the relation between physical events, but no *psycho-physical* laws governing their interaction. Suppose, however, that there is a harmony between mental and physical events. It might then be argued that in such a world it would be rational to assume that there are psycho-physical laws. If one did make that assumption one would consider that a physical event was necessary in the circumstances for the pain, which would be false. So once again, if there is room for rationality-judgments other than those based on the non-deductive inference from non-accidental generalisations discussed in Section Eight, then one has to exclude such rationality-judgments when giving an analysis of causation. However, even in this case, one can argue that there is no room for the (incorrect) assumption that there are psycho-physical laws. For P is not to modify his beliefs about the fundamentality of the psychic and the physical laws; yet if there are also psycho-physical laws, surely either the psychic or the physical laws cease to be as fundamental.

This argument in four steps shows that the conditionals required for my analysis of causation are indistinguishable from *restricted meta-inference conditionals*. A typical conditional occurring in the anlaysis of causation (*Nec 5*) would be:

> From the correct non-accidental generalisations and the assumption that all quasi-events occur which actually occur earlier than Y and which are distinct

existences from X and Y, and from the non-occurrence of X, one can infer (using deduction and the inferences discussed in Section Eight) that Y does not occur.

Note One: In the above conditional it is important that the non-accidental generalisations be referred to by the phrase 'the correct non-accidental generalisations' rather than explicitly listed. For if one supposed that the non-accidental generalisations were other than they actually are, then the inference would have to be made from the hypothetical non-accidental generalisations (in the analysis of a *hypothetical* case of cause and effect). For the same reason, the circumstances are referred to using the clause 'which actually occur' rather than explicitly listed.

Note Two: The conditionals 'If p, then q is highly likely' used in the Sufficiency Thesis can, I claim, also be treated as restricted meta-inference conditionals using, if necessary, probabilistic non-accidental generalisations. The kind of probability involved might be either *statistical* or *physical* depending on whether the non-accidental generalisation is of the form:

It is not an accident that a proportion α of quasi-events of kind K are followed by quasi-events of kind J

or of the form:

An event of kind K has a propensity α to be followed by an event of kind J.

In each case α is near 1, say 99%.

Section Ten: The Features on which Causation is Supervenient

The inferences discussed in Section Eight are, I claim, performed according to fixed rules, so whether the inference can be made depends only on the premises and the conclusion. Hence on this account of causation the features on which causation (relative to some given causal field) is supervenient are some or all of:

(1) The occurrence and distinct existences of X and Y

(2) The correct non-accidental generalisations and their degrees of fundamentality

(3) The quasi-events occurring no later than Y

I claim that given the causal field, whether X and Y are distinct existences is *supervenient* on which quasi-events X and Y are. So the features on which X causing Y is supervenient are contained in the non-accidental generalisations, their degrees of fundamentality, and the history of the world no later than X and Y.

Note One: The supervenience of the relation of being distinct existences is shown by the impossibility of conceiving of two pairs of events, identical in all respects except that the events of one pair but not of the other are distinct existences.

Note Two: Although this account of causation is in terms of non-accidental generalisations it is compatible with the claim that there is no irreducible non-accidental generality "in the objects". Thus I claim that the Conditional Analysis is compatible with the Regularity Account which I shall develop in the next chapter.

Note Three: The Conditional Analysis is also compatible with the claim that all non-accidental generalisations are, in fact, *universal* and hence laws of Nature. It is only within the antecedents of counterfactuals that one *has to* suppose that laws of Nature are broken (and hence not laws).

CHAPTER THREE

CAUSATION AND REGULARITIES

Introduction

This chapter is in two parts. In Part One I discuss what kind of regularity account one might expect to obtain and I reply in general terms to some objections which might be raised to any regularity account of causation. In Section One I discuss to what extent the Conditional Analysis might be used to support some kind of regularity account. In Section Two I consider what kind of regularity should be used in the Regularity Account. In Section Three I discuss the future-dependence of causation, to which a regularity theorist is committed, and I argue that the Regularity Account does not satisfy the Analysis-criterion, but merely the Reduction-criterion. In Section Four I discuss whether accidental regularities raise insuperable difficulties for a regularity account of causation.

Part Two is based on Chapter Three of Mackie's *The Cement of the Universe*. In Section Five I give a brief exposition of Mackie's account. I then propose, in Section Six, a reduction of causation to regularities, based on Mackie's account. Using the method of counter-examples, in Sections Seven to Ten I successively modify the reduction in order to cope with a wide range of proposed counter-examples.

As in Chapter Two, I ignore the possibility of backwards causation.

My aim in providing a regularity reduction of causation

is to examine *to what a metaphysician is committed* if, in the tradition of Hume, he claims that causation "in the objects" is nothing but the occurrence of various regularities, or that causation is *supervenient* on regularities. Only by considering a wide range of proposed counter-examples can one discover the extent of the commitment of a regularity theorist.

PART ONE

Section One: The Conditional Analysis and Regularities

One approach to a regularity account of causation would be to claim *as part of a metaphysical hypothesis* that non-accidental *generalisations* are systematically correlated with non-accidental *regularities*. One would then seek those features of non-accidental regularities on which their non-accidental character is *supervenient*. Such an approach would involve the following:

(1) A discussion of examples (actual or hypothetical) in which the non-accidental generalisation has no instance (or very few instances, and so, perhaps, does not count as a regularity. This situation would occur if there were examples of unique causation.

(2) A discussion of (actual or hypothetical) examples in which there are regularities which are not correlated with non-accidental generalisations. These I call accidental regularities.

(3) A discussion of some notion, such as the support of a regularity, which is used to provide an account of the degree of fundamentality of a non-accidental generalisation.

Although I do *not* adopt this approach, my discussion of unique causation,[2] accidental regularities,[3] and the support of a regularity,[4] could be adapted to the defence of a regularity account along these lines. Roughly, I argue that the notion of the support of a regularity enables one to permit some regularities with only one

---

1. By a regularity I mean a (universal or almost universal) generalisation with many instances. I later stipulate that in some cases a regularity may have only one instance.
2. See pp. 165,166.
3. See p.165 and pp.168,169.
4. See pp. 166-168.

instance (or with even no instances), and that, using this notion, one can distinguish (some) accidental from non-accidental regularities. The remaining examples of unique causation and accidental regularities can, I suggest, reasonably be rejected; this rejection is part of the regularity theorist's metaphysical hypothesis. Hence one could obtain an account of causation which satisfies the reduction-criterion but not the analysis-criterion. In this account, 'X causes Y' is *supervenient* on all the regularities and all the quasi-events occurring before the later of X and Y.

The approach I *do* adopt is to give an account of causation directly in terms of regularities, without considering the quasi-events occurring before the cause and effect (except as part of the distinct existences requirement). Such an approach has two advantages. The first is that if some successful objection is made to the Conditional Analysis the success of the Regularity Account is not jeopardised. The second advantage is that causal chains are not required, so the account is quite satisfactory even if one is considering the occurrence of point-events (rather than quasi-events, the location of which is a finite region of space-time).

Section Two: The Kind of Regularity being Considered

2.1 I claim that it is preferable to give a regularity account in terms of all the actual regularities, not merely those which are, at some time, observed or which are inferred from observations. Consider the following example. Suppose that there is an object with the mass of a whole galaxy but which is in size no greater than a pin-head. Suppose also that in this object events of kinds K and J occur, which never have been nor ever will be observed and the occurrence of which could not be inferred from observations. Finally,

suppose that events of kind *K* are regularly followed by events of kind *J*. Surely there is some causal connection between the *K* and the *J* (cause and effect or collateral effects). It would seem excessively ad hoc to deny the existence of such unobserved regularities as part of one's metaphysical hypothesis; therefore, the account should be based on the *actual* regularities. In addition, one might argue that there would be causation even if there were no observers (presumably, the regularity theorist is not committed to Phenomenalism).

2.2 I now argue in favour of permitting *time-dependent* regularities when analysing causation, from which it follows that the account of regularities given in Chapter One was oversimplified.[5] Consider the following hypothetical examples.

Example One: After 1700 A.D., and before 1100 A.D., whenever a certain kind of physical event (a P-event) has occurred in the brain, the person has felt pain. But no one ever felt pain between 1100 A.D. and 1700 A.D. Here I assume that it is coherent to speak of a person showing pain-behaviour (including the utterance of 'I am in pain') without being in pain. Furthermore, in the year 1990 a form of acupuncture will be discovered enabling people to easily control the occurrence of P-events. Pain will then be produced or prevented by producing or preventing P-events. I also assume that all methods for controlling pain work by preventing the occurrence of P-events.

In this case, it seems that after 1700 or before 1100 A.D. every P-event *caused* a pain. Yet the only available regularity for a regularity account is time-dependent.

Example Two: For half an hour every fourteen hours rubbing

---

5. See p. 45.

cinnibar on an opal is followed by the cinnibar turning into gold. This discovery is used to manufacture gold from cinnibar. It is to be assumed that there is no condition which happens to occur only in the half hour periods. In this case, surely one would say that rubbing the cinnibar on the opal caused it to turn into gold. Yet the only available regularity is time-dependent.

Now if these examples are coherent, the correlations might be described as time-dependent regularities. Yet, on the account of regularities given in Chapter One, there could be no time-dependent regularities. The question of what one calls these time-dependent correlations is not however important. What is important is whether they are to be considered as regularities for the purpose of the Regularity Account. If they are not considered to be regularities, then the regularity theorist would be committed to denying that there is any causal relation. This seems odd. It would not be excessively ad hoc simply to deny *as part of one's metaphysical hypothesis*, that time-dependent regularities do occur; nonetheless, it seems better to avoid an unnecessary metaphysical hypothesis, using if required time-dependent regularities in the account of causation.

As in Chapter One, I stipulate that quasi-events are to be described in such a way that any change made in their position in space-time would result in the description of *some* possible quasi-event. However I now permit time-dependent regularities. For instance I permit regularities which are restricted to the union of some suitable, simply described, sequence of temporal intervals.

Note: In this case, some rule is required such as the rule that the more complicated is the sequence of temporal intervals the larger is the number of instances of the correlation which is required for it to be considered a regularity.

2.3 In Section 2.2 I have permitted, in the account of causation, the use of regularities which are spatio-temporally restricted. It might be suggested that one should always use spatio-temporally restricted regularities. There is, however, no obvious advantage in so doing, and I now point out some difficulties with three restrictions that might be proposed.

Proposal One: The regularities are to be spatially restricted to events occurring at or near the earth's surface.

Discussion: The obvious difficulty with this proposal is that one should not deny that there could be causal relations between events which are quite unlike those which occur on the Earth (see the example in Section 2.1).

Proposal Two: The regularities are to be temporally restricted to those occurring no later than the cause and the effect.

Discussion: The difficulty with this proposal is that in some cases the regularity theorist is committed to the future-dependence of causation (discussed more fully in Section Three). For example, suppose that there are two simultaneous initial events A and B followed by a third event C. Suppose also that, throughout the history of the universe, events like A are regularly followed by events like C and events like C are regularly preceded by events like A. Suppose also that events like C are *not* regularly preceded by events like B and that events like B are *not* regularly followed by events like C. Then, surely, A causes C but B does not cause C (nor do A and B overdetermine C). So, in this case, the truth of 'A causes C' is, on the Regularity Account, dependent on events occurring after A and C. The regularity theorist could avoid the future-dependence of causation only at the cost of *either* the denial that there are initial events

or the denial that there is a causal relation between *initial* events. In either case Proposal Two commits the regularity theorist to the inclusion in his metaphysical hypothesis of an unnecessary and ad hoc stipulation.

Proposal Three: The regularities are to be restricted to those holding earlier than the utterance (or thought) of 'X causes Y'.

Discussion: The difficulty with this proposal is that it makes the truth of causal sentences depend on their utterance at some time or other. But it would be unnecessary and ad hoc if the regularity theorist were to exclude causal relations which have not been and which will not be discovered.

2.4 Finally, I claim that the regularities need not correspond to universal generalisations; it suffices that there are few exceptions (other than, perhaps, those due to *immanent* causation). The purpose of this proviso is to avoid unnecessary commitments on the part of the regularity theorist. However, for the sake of simplicity, I usually assume that the regularities do correspond to *universal* generalisations.

Section Three: The Future-dependence of Causation

In this section, first I examine the commitment of the regularity theorist to the future-dependence of causation. Then I show that the Regularity Account can be expected to satisfy only the Reduction-criterion, not the Analysis-criterion.

3.1 The Future-dependence of Causation

The regularity theorist is committed to the thesis that the truth of some causal sentences depends on the occurrence of events

later than both the cause and effect.[6] Consider the following two arguments.

Argument One: If an initial event X causes an event Y, the required regularity (events like X are followed by events like Y) depends on events occurring later than both X and Y. An example of this situation was discussed in Section 2.3.

Argument Two: Suppose that a certain regularity holds up to, say, the year 2,000 A.D. Whether or not it is accidental might depend on events occurring after the year 2,000 A.D. An example of this situation is as follows. Suppose that this universe is spatially finite and that it came into existence 10,000,000,000 years ago and will cease to exist in the year 2,000 A.D. Suppose that throughout this universe (sufficiently) heated water always boils. Presumably, the regularity theorist would consider this to be a *non-accidental* regularity and would say that heating water *causes* it to boil. Now consider another universe exactly like the first for 10,000,000,000 years but which continues to exist for another 10,000,000,000 years after the year 2,000 A.D. In the second universe the regularity 'Sufficiently heated water boils' only holds *up to the year 2,000 A.D.* After 2,000 A.D. there is an extremely *low*, but positive (say one in 1,000,000,000), frequency of cases in which heated water *boils*, and there is an extremely *high* frequency of cases in which heated water *freezes*. In the second universe one would say that there was a non-accidental *time-dependent* regularity that heated water boils for the first 10,000,000,000 years. Consider yet a third universe which is exactly like the first and second universes up to the year 2,000 A.D. but which continues indefinitely after the year 2,000 A.D. In the third universe (as in the second) after the year 2,000 A.D. there is an extremely *low*, but positive (one in 1,000,000,000), frequency of cases in which heated

6. This point is due to W.D. Joske.

water boils. *Relative* to the knowledge that there is a one in 1,000,000,000 frequency of cases in which heated water boils, it is highly probable that *eventually* there is some interval of 10,000,000,000 years throughout which, by chance, heated water *boils*. I suggest that in the third universe one should not say that there is a non-accidental time-dependent regularity that heated water boils for the first 10,000,000,000 years. Rather, one would say that it was an *accident* that for 10,000,000,000 years heated water boils. The comparison of these three universes shows that whether a regularity is accidental or not might depend on what happens in the distant future. Therefore, the truth of 'Heating the water caused it to boil (in the year 1978)' depends on what happens billions of years hence.

3.2 Someone might object to the Regularity Account on the grounds that the future-dependence of causation is *Counter-intuitive*. In this subsection I reply to this objection.

First I define hypothetical and categorical future-dependence as follows.

Definition: A sentence p apparently about events X and Y (for example, 'X causes Y') is said to be *hypothetically* future-dependent if there is some true sentence q which is about events later than X and Y, such that the counterfactual 'If ~q, ~p' is correct.

Note: I do not assume that the above counterfactual is necessarily one of the kinds discussed in Chapters One and Two.

Definition: A sentence p is said to be *categorically* future-dependent if the truth or falsity is not fixed (determined) at the time of the events which p is apparently about.

Note: A sentence may be both hypothetically and categorically future-dependent.

Example One: If it is now fixed that no atom bomb will be used again in warfare, then the sentence 'The last atom bomb ever to be used in warfare was dropped on Nagasaki' is hypothetically but not categorically future-dependent.

Example Two: Suppose that backwards causation is coherent and that a later event X causes an earlier event Y. The occurrence of Y - like any other event - is fixed no later than the time of its occurrence. Hence the sentence 'Y occurs' is not *categorically* future-dependent. However, the sentence 'Y occurs' is *hypothetically* future-dependent for X causes Y and so the counterfactual 'If X will not occur, Y does not occur' is correct.

Example Three: Suppose that the occurrence of some event X is *absolutely* sufficient in the circumstances for the occurrence of some event of kind *K* - not merely probabilistically sufficient. Suppose that one treats the sentence 'X will cause some event of kind *K*' as a sentence about X rather than about events of kind *K*. The counterfactual 'If no event of kind *K* will occur then X will not cause an event of kind *K*' is correct (the occurrence of the effect is a necessary condition for there to be a causal relation). Hence the sentence 'X will cause an event of kind *K*' is *hypothetically* future-dependent. But this sentence is not *categorically* future-dependent because the occurrence of X is absolutely sufficient in the circumstances for the occurrence of some event of kind *K*.

I now claim that only the *categorical* future-dependence of causation would be sufficiently counter-intuitive to warrant the rejection of the Regularity Account. First notice that on any version

of the Regularity Account the truth of 'X causes Y' depends on there being a regularity 'Events like X are followed by events like Y', and hence depends on events which are not in the *spatio-temporal vicinity* of X and Y. I suggest that the hypothetical future-dependence of the sentence 'X causes Y' is merely one aspect of this dependence of causal sentences on events not occurring in the spatio-temporal vicinity of the cause and effect. Therefore, the hypothetical future-dependence of the sentence 'X causes Y' is no more peculiar than the dependence of that sentence on events occurring earlier than X and Y, or on events occurring some distance away from X and Y. It follows that the intuition that causation is not future-dependent, unless it is accompanied by an intuition that causation is only dependent on events in the spatio-temporal vicinity of the cause and effect, is an intuition that causation is not *categorically* future-dependent.

It remains to show that the future-dependence of causation is not *categorical*. If one could assume that *Determinism* is correct one could argue as follows:

> If, on the Regularity Account, X causes Y then there is a regularity such as 'All (quasi-)events qualitatively identical to X are followed by (quasi-)events qualitatively identical to Y'. This regularity is fixed at the time of X, since all events are fixed at the time of X.

Notice that the regularity theorist is assuming a hypothesis stated for the sake of simplicity in causal terms. Strictly speaking, the regularity theorist is assuming that the regularity reduction of Determinism is correct.

However, the regularity theorist need only assume that it is *fixed* (either eternally or by the initial state of the universe) that all those regularities which, on the Regularity Account, are non-

accidental have enough instances to be deemed regularities. For it is a consequence of the Regularity Account itself that the generalisations which correspond to those regularities deemed non-accidental are non-accidental generalisations (that is, roughly speaking, laws of Nature) and hence are eternally fixed. So the occurrence of a regularity is fixed as soon as it is fixed that the generalisation has enough instances. Again, this account involves the notion of fixity which is a causal notion. Thus the account should be replaced by its rather complicated regularity reduction.

Note One: I later permit regularities with only one instance provided they are supported by regularities with many instances.[7] The supporting regularities need only concern sub-microscopic events. Consequently, the regularity theorist only has to claim that it is eternally fixed that the universe exists long enough for every "fundamental law", relating sub-microscopic instances, to have many instances.

Note Two: The dependence of causation on events occurring outside the spatio-temporal vicinity of cause and effect is compatible with there being *marks* of causation which can be observed in the *vicinity* of the cause and the effect (for example, one could use the Ducasse criterion). The regularity theorist might give an explanation of the reliability of the Ducasse criterion as follows. Rational animals could only have evolved from animals of fairly low intelligence which nonetheless display purposive behaviour. It is plausible that such animals could not display purposive behaviour unless they could *spot* causal situations. Hence rational creatures could only evolve in an environment where the Ducasse criterion - or a similar criterion - was fairly reliable.

---

7. See p. 166.

3.3 The Criterion which the Regularity Account Satisfies

The reply in Section 3.2 to the objection that the future-dependence of causation is counter-intuitive was based on the claim that it is fixed that the universe exists for a sufficient length of time. I now use this observation to show that the Regularity Account does not satisfy the Analysis-criterion.[8] The metaphysical theory that there are irreducibly nomic universal generalisations is, I suggest, a coherent one. On that theory there could be two universes in each of which an initial event X is followed by another event Y and in each of which the universe then ceases to exist. In one of these universes it is *nomically necessary* that X is followed by Y; in the other it is *not*. So in one of the universes X *causes* Y, in the other universe X *does not cause* Y. But the regularity theorist cannot distinguish between the two universes. Hence there are two conceivable situations in which the reductianda differ in truth-values but the regularity reductians have the same truth-value. Therefore, the Regularity Account does not satisfy the Analysis-criterion.

Section Four: Accidental Regularities

Accidental regularities are widely held to raise special difficulties for a regularity account of causation. Thus Mackie says: "The problem ... of distinguishing causal from accidental regularities is the great difficulty for any regularity theory of causation."[9] It seems appropriate to discuss in fairly general terms two objections based on the possibility of accidental regularities, before considering the details of the Regularity Account of causation.

---

8. A similar argument could be based on in principle undiscoverable accidental regularities, see p.140.

9. Mackie, J.L., *The Cement of the Universe*, p.196.

4.1 One objection is that if a regularity account is correct there is no *real*, that is mind-independent, difference between non-accidental and accidental regularities; laws of Nature are cosmic accidents. I do not consider this objection to be a weighty one; if one has a metaphysical hypothesis on which there is no irreducibly causal element "in the objects" then it is not surprising that causation does not appear in an account of the mind-independent difference between causal and accidental regularities. There is, of course, a difference between those regularities which one *calls* causal or non-accidental and those which one *calls* accidental, but this difference need not be described *in causal terms*.

Related to this objection is a weightier one to which I now reply, namely the objection that the regularity theorist is not able to give an account of the differences between accidental and non-accidental regularities, without resorting to nomic necessity or some similar notion. I consider three kinds of accidental regularity; in each case I show how, if at all, the regularity theorist distinguishes them from non-accidental regularities.

4.2 Spurious Accidental Regularities

Some regularities are not accidental but might nonetheless loosely be called accidental. For example, Taylor considers a thousand specially decorated matches all of which light when struck.[10] Now suppose someone, who wanted to disprove a regularity theory, ensured that all the matches ignited when struck. In that case the regularity is not accidental but an example of the *collateral effects* of an agent's activity. The agent could be replaced by a machine, in which case the regularity becomes a standard case of collateral effects.

---

10. Taylor, R., *Action and Purpose*, pp. 24, 25.

Another example of a spurious accidental regularity which is, in fact, an example of collateral effects is the correlation between the rising of the star Sirius above the horizon and the flooding of the Nile. This regularity is explicable in terms of the time taken for the flood to move down the Nile and the regularity of the wet season in Ethiopia.

The regularity-theorist has to show how collateral effect situations are distinguished from situations of cause and effect. This will be discussed in Sections 7.4 and 10.4.

4.3 Accidental Regularities that could not even in Principle be Discovered

Consider again the example of the three universes (Section 3.1). The realist about nomic necessity could consider a fourth universe which, like the first, ceases to exist after 10,000,000,000 years but in which it is a *law* that there is a very *low* but positive probability that heated water boils. By chance, in this universe heated water always boils. In this case the regularity 'Whenever water is sufficiently heated it boils' is accidental, but it could not, even in principle, be discovered to be accidental.

The regularity theorist can, I claim, simply deny that there are any such in principle undiscoverable accidental regularities. For it is only by supposing that there are laws of Nature of a probabilistic kind that one can ensure the coherence of the description of a universe in which there are in principle undiscoverable accidental regularities. The regularity theorist would deny that there is any difference between the fourth universe and the first universe.

4.4 In Principle Discoverable Accidental Regularities

Consider the following example of Mackie's:

> "Suppose that we manufacture a number of atomic bombs ... each of which will explode if and only if a nuclear disintegration of a certain kind A occurs spontaneously before a nuclear disintegration of some other kind B occurs within the same core ... It is a matter of pure chance which [bombs] do [explode] ... it just happens that a red spot has been painted on all and only the bombs that in fact explode." [11]

The regularity theorist's account would be weakened if his metaphysical hypothesis contained the ad hoc assertion that accidental regularities like the one in this example do not occur. For, given some random device used only moderately often, it is quite likely that one of the many features present in some cases but not in others is in fact correlated with one of the results of using the random device. For example, consider an experiment designed to test for precognition which is used only ten times and in which there is a random number generator. It is quite likely that some fairly simple feature of the environment happens to be correlated with a particular output of the random number generator, say the occurrence of some particular frequencies in the vibrations of the air and the floor of the laboratory. In this case the objection to the Regularity Account is that there is a difference, which is in principle discoverable, between accidental and non-accidental regularities, but this difference cannot be described without resorting to the notion of nomic necessity (or a related notion). I now consider four replies to this objection, of which only the last two are satisfactory.

Reply One: The regularity-theorist could insist that regularities should have *very many instances*. So, although it is conceivable

---

11. Mackie, J.L., *The Cement of the Universe*, p. 198.

that there be accidental regularities, the chance of their occurring is *extremely small*. The regularity theorist would then claim that *in fact* there are no such accidental regularities. Thus, in the above example, if the machine is used 1,000 times, the chance of an accidental correlation with simple features of the environment is extremely small.

There are two objections to this reply. The *first* is that there could still be cases of accidental *overdetermining* regularities. For instance, consider Taylor's example of the decorated matches; the accidental regularity that all decorated matches light seems to support the claim that the decoration and the dryness *overdetermine* the matches lighting. Such accidental regularities could easily occur since most matches are dry and consequently light when struck. This objection can be handled in various ways; one might require that if there is a regularity 'All Ys are preceded by Xs' then 'All Ys are preceded by Xs or Zs' is deemed not to be a regularity. (This is a consequence of the selection rule stated below in Section Eight.[12]) *The second objection* is that the universe is very large (perhaps infinite) so accidental regularities with many instances probably occur somewhere. For this reason I find the first reply unsatisfactory.

Reply Two: The second reply is that, as part of one's metaphysical hypothesis, one can assert that Determinism is correct, and also that there is a single initial event. In this way, it might be argued, all accidental regularities are cases of *collateral* effects.

This reply is also unsatisfactory. Determinism is itself a *causal* notion and therefore the regularity theorist should not seek to distinguish between accidental and non-accidental (causal) regu-

12. See p.164.

larities in terms of a causal notion. Suppose that events of kind A have always been followed by events of kind B. Does the occurrence of an A *determine* the occurrence of a B? Not if the regularity is *accidental*. So until the regularity theorist has distinguished accidental from non-accidental regularities, he cannot give an account of the notion of Determinism.

Reply Three: I suggest that there are rather general statistical regularities of the kind 'Quasi-events of kind *K* are *statistically independent* of quasi-events of kind *J*'. (An account of statistical independence will be given later.[13]) A statistical regularity of this kind with very many instances outweighs an ordinary regularity with few fewer instances. Thus there is a statistical regularity that nuclear processes are not influenced by macroscopic events of a rather vaguely defined kind, which includes putting spots of paint on bombs.

Reply Four: (considered in greater detail later[14]) There are *fundamental* regularities characterised by their *simplicity* when stated in scientific realist terms (that is, involving sub-microscopic events). A regularity is *accidental* unless it is derived from fundamental regularities.

Two conclusions can be drawn from the discussion of accidental regularities in this section. On the one hand, the regularity theorist can give a satisfactory account of the distinction between in principle discoverable accidental regularities and non-accidental regularities. On the other hand, there is no need for the regularity theorist to distinguish between non-accidental regularities and accidental regularities that are not even in principle discoverable.

---

13. See p.168.

14. See p.169.

PART TWO

Section Five: A Brief Exposition of Mackie's Discussion of Causal Regularities in Chapter Three of *The Cement of the Universe*.

Because my regularity account of causation is based on Mackie's account of causal regularities, I now briefly expound Mackie's account.

5.1 Before discussing to what extent one can give a regularity account of causation, Mackie expounds Mill's account of *complex regularities*.

In Chapter Three of *The Cement of the Universe*, Mackie uses the terms 'necessary condition' and 'sufficient condition' in an *extensional* sense, as follows:

> 'X is a necessary condition for Y' means 'Whenever an event of type Y occurs, an event of type X also occurs';
>
> and 'X is a sufficient condition for Y' means 'Whenever an event of type X occurs, an event of type Y also occurs'.[15]

Given any type of event or situation P one can seek *factors*, that is types of events or situations A, B, C etc. such that whenever all of A, B, C etc. occur then a P follows, but none of the A, B, C etc. are redundant. Thus if the occurrence of an A and a B and a C etc. is denoted by 'ABC...' and A, B, C etc. are factors for P, then ABC... is sufficient for P and earlier than P. Perhaps there are other sufficient conditions for a P, such as DGH... and JKL.... If it happens that P are always preceded by ABC... or by DGH... or by JKL... or etc...., then one obtains the complex regularity:

> An (ABC... or DGH... or JKL... or...) is necessary and sufficient for P and followed by a P.

---

15. Mackie, J.L., *The Cement of the Universe*, p.62.

For the sake of simplicity Mackie considers the example of a regularity:

> An (ABC or DGH or JKL) is necessary and sufficient for a P and is followed by a P.

ABC is then a *minimal sufficient condition* in the sense that no part of ABC is also sufficient for P. A is, in general, an *insufficient* but *non-redundant* part of ABC which is, in turn, an *unnecessary* but *sufficient* condition for a P. Mackie calls A an *inus* condition for P.[16] Some of the factors might be the *absence* of events; for example, not taking an antidote could be an inus condition for death. If C is a type of event, let '$\overline{C}$' denote the absence of a C. The typical complex regularity is:

> All $AB\overline{C}$ or $DG\overline{H}$ or $JK\overline{L}$ are followed by some P and all P are preceded by some $AB\overline{C}$ or $DG\overline{H}$ or $JK\overline{L}$.

As in Chapter Two, Mackie notes that factors are relative to a causal field. Now, ($AB\overline{C}$ or $DG\overline{H}$ or $JK\overline{L}$) is in *disjunctive normal form*. In general, any regularity between P and a truth-functional compound of the occurrence of earlier factors can be put into the form considered by Mackie.

Mackie defends the use of complex regularities; here he follows Mill rather than Hume. His argument is that one often does not *know* the complete regularity. In that case if it is required that all P are preceded by Q for Q to be a cause of P, then one could not know the cause of P. Yet in many cases one claims to know some causes of a given phenomena - at least in a weak sense of 'know'. If the cause "in the objects" of the P is an *inus* condition for the P, then one does indeed know *some* of the causes of the P. Mackie

---

16. *Ibid.*, p.62.

argues that knowledge of some inus conditions, even when combined with ignorance of other inus conditions, provides one with the theoretical knowledge which a knowledge of causes *should* provide. Thus Mackie says:

> Knowing that something of the form (AX or Y) - where A is known, but X and Y are not - is both necessary and sufficient for P in F, we may well have reason to believe that X, whatever it may be, is often present; if so, we can infer from an observed occurrence of A that P is fairly likely to follow. [17]

For example, X might be the failure to take an antidote and A might be the event that the poison is taken. Even if one does not know what antidotes there are, one could still predict that if the poison is taken the person is likely to die. Similarly, if causes are inus conditions, one can infer causes from effects, the absence of causes from the absence of effects, and the absence of effects from the absence of causes. All such inferences involve judgments that other factors are unlikely to occur - without necessarily knowing precisely what these other factors are. Such judgments are based on the consideration that factors which are likely to occur are, in general, likely to be discovered using the Method of Differences. Mackie discusses *Mill's methods* at some length.[18] He argues that the assumption that there is a complex regularity - together with various observations - enables one, using Mill's methods, to discover at least some inus conditions (the more frequently occurring ones). Moreover, knowledge of inus conditions has the *practical* use which one expects knowledge of causes to have. For if A is an inus condition for P and one produces an A, one can *hope* that the other parts B, C etc. of the minimal sufficient condition for P are present, and so one

17. Mackie, J.L., *The Cement of the Universe*, p.67.

18. *Ibid.*, pp. 68-74.

can also *hope* that a P will occur. Similarly, by preventing an A one has done *something* to prevent a P, even though there might be other minimal sufficient conditions for P.

5.2 Thus in the first part of Chapter Three Mackie has argued that a regularity theory based on complex regularities has considerable merit. He then argues, however, that singular causal statements differ in *meaning* from the assertion that some sequence is an instance of a regularity.[19] Mackie's argument is that one can make judgments about causation (especially in mental examples) without being committed to any generalisation. Here I think one should distinguish between being committed to the occurrence of some regularity and being able to provide the regularity. I claim that causal statements are *universalisable*, that is, if someone asserts that X causes Y he is committed to there being some kinds of event K and J and such that

(1) All K cause J in circumstances like those in which X causes Y.

(2) K, J and the description of the circumstances do not involve causal notions.

and (3) K and J are not described in such a way as to severely limit the possible number of occurrences of K and J.

The universalisability of causation can be compared to universalisation in Ethics. Although I could judge that it is wrong to say what I did without being able to give a more general rule, I am committed to the existence of some (non-ethical) feature such that in any situation containing these features one person is wrong to say what I said to another person. Thus I do not find Mackie's argument entirely convincing. However, I have argued (see Section 3.3)

19. Mackie, J.L., *The Cement of the Universe*, p.77.

that no regularity account could satisfy the analysis-criterion; accordingly I agree with Mackie when he goes on to say:

> The crucial and outstanding question is to what extent such complex regularities as we have described constitute causation as it is in the objects. [20]

Here Mackie seems to be considering what he calls the *factual analysis* of causation in terms of regularities. Mackie argues that regularities are *not the whole* of what constitutes causation "in the objects". His argument is that if types of event A and B are *collateral* effects of type of event C, and the A are always followed by the B, then A is an *inus* condition for B, but the A do not *cause* the B. Mackie's example (based on Broad's example of people going to work) is that in which A is the sounding of factory hooters in Manchester and B is the event that London workers leave their work. C is the event that it is 5 o'clock in England.[21] A are regularly followed by B, but the A do not cause the B. Mackie considers Mill's proposed solution to this problem, namely that one knows of conditions in which A would not be followed by B and so the regularity 'All A are followed by B' is not *counterfactually* unconditional. The argument which Mackie uses against Mill is as follows:

If A and B are collateral effects of C, there are complex regularities:

(1) All (CX or Y) are followed by A and all A are preceded by CX or Y

and (2) All (CZ or W) are followed by B and all B are preceded by (CZ or W).

In this case X, Y, Z and W could be very complex factors and in fact Y and W might never actually occur. Mackie then argues that all $A\overline{Y}$ are preceded by C and all CZ are followed by B, hence all $A\overline{Y}Z$ are

---

20. Mackie, J.L., *The Cement of the Universe*, p.80.

21. *Ibid.*, p.81.

followed by B. Furthermore, the regularity 'All $A\overline{Y}Z$ are followed by B' is *counterfactually* unconditional. So, even leaving aside the problem of providing a regularity account of what it is to be *counterfactually* unconditional, Mill's solution seems to fail because in cases of collateral effects A is not causally sufficient for B. Mackie concludes that one requires the notion of *causal priority* even if one is considering, not the meaning of causation, but what causation is "in the objects". Here I disagree with Mackie - one of my aims in providing a regularity account is to avoid any appeal to *causal priority*.

Section Six: A Reduction based on Mackie's Account of Causal Regularities

In spite of Mackie's denial that the Regularity Account is a complete account of causation "in the objects",[22] I develop an account in which causes are reduced to *inus* conditions (or to factors which are *at least* inus conditions). I begin with some notation.

6.1 Notation

(1) Quasi-events are denoted by capital letters
(No distinction is made between the use of A, B, C etc. and X, Y, Z etc.)

(2) If X is a quasi-event, $\overline{X}$ and ~X both denote the *absence* of X. But ~X is only used where the absence of X is itself a quasi-event. For example, if someone failed to take an antidote which no one knew was an antidote, the failure to take it would not be a quasi-event, but merely the absence of one. But if the antidote were routinely taken as a precaution, the failure to take the antidote would be considered a quasi-event.

---

22. Mackie, J.L., *The Cement of the Universe*, p.86.

(3) Square brackets are used to convert a reference 'X' to a particular quasi-event to a predicate [X] interpreted as 'qualitatively identical to X'.

(4) Capital letters in *italics* denote *kinds* of quasi-events.

(5) Usually, no mention is made of the *causal field* to which both causation and regularities are relative.

(6) The regularity[23]:

All [$AB\overline{C}$] or [$DG\overline{H}$] or [$JK\overline{L}$] are followed by [P], and
all [P] are preceded by [$AB\overline{C}$] or [$DG\overline{H}$] or [$JK\overline{L}$]:

is written as:

Reg: [$AB\overline{C}$] $\vee$ [$DG\overline{H}$] $\vee$ [$JK\overline{L}$] $\leftrightarrow$ [P].

This notation is also used if [P] is simultaneous with the latest of [A], [B], etc.

(7) Suppose that X, Y and Z but neither a [W] nor a [V] occur on some occasion and there is a regularity:

Reg: [$XZ\overline{W}$] $\vee$ [V] $\leftrightarrow$ [Y]

In this case, X would be considered both necessary and sufficient in the circumstances for Y ($Z\overline{W}$ is the circumstances). In order to state the Necessity Thesis in terms of regularities I require other arrows in addition to '$\leftrightarrow$'.

Reg: [$XZ\overline{W}$] $\vee$ ... $\leftarrow$ [Y]

is interpreted as:

All [Y] are *preceded* by (or simultaneous with) [$XZ\overline{W}$] or ...

Again,

Reg: [$XZ\overline{W}$] $\vee$ ... $\rightharpoonup$ [Y]

is interpreted as:

The *vast majority* of [$XZ\overline{W}$] or ...
are followed by (or simultaneous with) [Y].

---

23. I refer to regularities by the corresponding generalisations.

If both

Reg: $[XZ\overline{W}] \vee \ldots \leftarrow [Y]$

and

Reg: $[XZ\overline{W}] \vee \ldots \rightarrow [Y]$

occur then I write:

Reg: $[XZ\overline{W}] \vee \ldots \leftrightarrow [Y]$.

(8) I also require, when giving an account of causation in terms of regularities, what I call *particular presentations of regularities* (PPRs). A PPR is a particular sequence of quasi-events considered as an instance of a regularity. For example, there is said to occur (or exist) the PPR,

Pres: A, P; $[A] \vee [B] \leftrightarrow [P]$

if and only if:

(1) There is a regularity

Reg: $[A] \vee [B] \leftrightarrow [P]$

and (2) A and P but not B occur.

In the same situation, there would *also* be the PPR,

Pres: $A \vee B$, P; $[A \vee B] \leftrightarrow [P]$

In general, the regularity of the PPR is always put *either* in the form:

[ ] arrow [ ]

where the contents of both square brackets occur (and are listed to the right of 'Pres') *or* in the form:

[ ] $\vee$ [ ] arrow [ ],

where the contents of the first and third square-brackets occur (and are listed to the right of 'Pres'), but the contents of the second bracket do not occur. Any arrow may be used in the notation for PPRs.

Note: Thus the square brackets are not only used to convert the

reference X to the qualitative identity predicate [X], but square brackets are also used to group together disjuncts in the regularity when describing the PPR.

6.2 Some Preliminary Remarks on the Regularity Account

(1) There are cases of symmetric overdetermination in which the regularity is, say,

Reg: $[AB\overline{C}] \vee [DG\overline{H}] \vee [JK\overline{L}] \leftrightarrow [P]$

and A, B, D, G and P but neither C, H, J nor K occur. In that case, if one holds to the necessity of the Necessity Thesis, A does *not* cause P, yet [A] is an inus condition for [P]. In such cases there is *not* the PPR:

Pres: $AB\overline{C}$, P; $[AB\overline{C}] \vee [DG\overline{H} \vee JK\overline{L}] \leftarrow [P]$,

since $DG\overline{H} \vee JK\overline{L}$ occurs. But there is the PPR:

Pres: $AB\overline{C} \vee DG\overline{H}$, P; $[AB\overline{C} \vee DG\overline{H}] \vee [JK\overline{L}] \leftarrow [P]$.

hence I require in the 'A causes P' that the PPR be considered not merely the regularity. Notice that this requirement is harmless if one accepts the sufficiency of the Sufficiency Thesis. For if there is the PPR:

Pres: $AB\overline{C} \vee DG\overline{H}$, P; $[AB\overline{C} \vee DG\overline{H}] \vee [JK\overline{L}] \rightarrow [P]$

there is also the PPR:

Pres: $AB\overline{C}$, P; $[AB\overline{C}] \rightarrow [P]$

(2) There is no need to mention the *absence* of a quasi-event unless that absence is itself a quasi-event. *On the one hand,* if $\overline{C}$ is not a quasi-event, it is not a cause. *On the other hand,* if there is a regularity Reg $[AB\overline{C}] \vee \ldots \leftrightarrow [P]$, then *either* C occurs in very few cases, in which case there is also the regularity Reg $[AB] \vee \ldots \leftrightarrow [P]$, *or* $\overline{C}$ is itself a quasi-event and '$\overline{C}$' could be rewritten as '~C'.

(3) Conjunctions of quasi-events are quasi-events, so by (2), [A] is an inus condition for [P] if and only if there is some X such that there is a regularity: Reg: [AX] ∨ ... ↔ [P].

Moreover, the disjunction of two quasi-events is a quasi-event unless it is part of every normal state-of-affairs. But in the latter case it should not be mentioned in the regularity. So if [A] is an inus condition for [P] there is a regularity:

Reg: [AX] ∨ [Y] ↔ [P] [24]

(4) Where there is a regularity

Reg: [AX] ↔ [P]

or Reg: [A] ∨ [Y] ↔ [P]

or Reg: [A] ↔ [P],

[A] is not, strictly speaking, an *inus* condition for [P]. [A] is what Mackie calls 'at least an inus condition' for [P]. The account of causation is in terms of what are at least inus conditions. Moreover if the Necessity Thesis is deemed sufficient, one would consider regularities such as:

Reg: [AX] ∨ [Y] ← [P],

where [A] is not part of a *sufficient* condition for [P] but rather is part of an *as sufficient as possible* condition for [P]. For example, in the case of Mackie's chocolate machine L, putting the shilling in is an insufficient but necessary part of a condition which, although not sufficient for obtaining the chocolate bar, is *as sufficient as possible*.

(5) The regularity,

Reg: [AX] ∨ [Y] ↔ [P]

can be written,

---

24. Mackie himself uses this convenient form. See *The Cement of the Universe*, p.71.

Reg: $[(AX \vee Z)(AX \vee \overline{Z})] \vee [Y] \leftrightarrow [P]$.

for any Z, yet, presumably, [AX ∨ Z] is not an inus condition for [P]. Some restriction is required on the kind of part which A may be of the "unnecessary but sufficient" condition AX. As Kim points out,[25] if one simply requires that the regularity be put into disjunctive normal form, then one seems to make causal relations depend on the description of the (quasi-)events. For instance, in Section 7.1 I give an example in which a quasi-event D could be described as a disjunction, namely ~C ∨ A.[26] I suggest that *either* A is to be the same as AX *or* A and X are both to be *regular* parts of AX.[27]

6.3 Taking these five remarks into consideration, I propose the following formulation of "Mackie's Regularity Account".

Reg 1: There are three cases:

Case One: If both the Necessity Thesis and the Sufficiency Thesis are necessary, then 'A causes B' is reduced to:

(1) B is a distinct existence from the conjunction of A and all earlier quasi-events.

and (2) One of the following PPRs occurs:

Pres: C, B; [C] ↔ [B]

Pres: C, B; [C] ∨ [D] ↔ [B],

For some quasi-events C and D, where *either* A is the same as C *or*, for some X, AX is the same as C and both A and X are regular parts of C.

Case Two: If the Necessity Thesis is both necessary and sufficient

---

25. Kim, J., "Causes and Events: Mackie on Causation", *Journal of Philosophy*, vol.68(July 1971), pp. 433-435.

26. See p.157.

27. For the definition of a *regular* part of a quasi-event, see p. 43.

then 'A causes B' is reduced to:

(1) B is a distinct existence from the conjunction of A and all earlier quasi-events.

(2) One of the following PPRs occurs:

Pres: C, B; [C] ← [B]

Pres: C, B; [C] ∨ [D] ← [B],

for some quasi-events C and D, where *either* A is the same as C *or*, for some X, AX is the same as C and both A and X are regular parts of C.

and (3) There is *no* PPR

Pres: C*, B; [C*] ∨ [D*] ← [B]

such that C is a part of C* (regular or otherwise) and such that [C*] are *followed* by (or simultaneous with) [B] in a significantly greater proportion of cases than [C] are followed by (or simultaneous with) [B].

Case Three: If the Sufficiency Thesis is both necessary and sufficient, then 'A causes B' is reduced to

(1) B is a distinct existence from the conjunction of A and all earlier quasi-events.

(2) There is a PPR,

Pres: C, B; [C] → [B]

where *either* A is the same as C *or*, for some X, AX is the same as C and both A and X are regular parts of C.

and (3) There is *no* PPR

Pres: C*, B; [C*] → [B]

such that C* is a part of C (regular or otherwise) and such that [B] are *preceded* by (or simultaneous with) [C*] in a significantly greater proportion of cases than [B] are *preceded* by (or simultaneous with) [C].

Note One: The distinct existences requirement is that B is a distinct existence from the conjunction of A and all earlier quasi-events. This requirement is necessary because of the example of Joanne giving birth and Harry becoming a father[28] and the example of the amoeba Jojim dividing into two amoebae Jo and Jim.[29]

Note Two: Condition (3) is required in Case Two, to ensure that the complete cause C is as sufficient as possible. Likewise, condition (3) is required in Case Three to ensure that the complete cause C is as necessary as possible. Suppose there is a PPR,

Pres: C, B; [C] ∨ [D] ↔ [B],

and very rarely, by accident, when a [C] occurs and [E] also occurs, but on no other occasion has an [E] occurred. If on the occasion in question E occurred, then one would also have the PPRs

Pres: C ∨ E, B; [C ∨ E] ∨ [D] ↔ [B],

and Pres: CE, B; [CE] ∨ [D] ↔ [B].

Yet E is *causally irrelevant* to B. The above PPRs are excluded by condition (3) in Case Two and Case Three.

Note Three: I concentrate on the simplest case (Case One), where the Necessity Thesis and the Sufficiency Thesis are both required.

Note Four: The occurrence of the regularity Reg: [A] ↠ [C] is *not* an automatic consequence of the occurrence of Reg: [A] ↠ [B] and Reg: [B] ↠ [C]. So I *stipulate* that if A to B to C is to be called a *causal chain* then A causes C, that is, where causation is not transitive one should not speak of causal chains.

---

28. See p.44.

29. See p.87.

Section Seven: Some Objections to "Mackie's Regularity Account"

I now discuss some objections to the account *Reg 1* that might be raised.

7.1 Objection One: (Due to Kim)

> Assume AB is a minimal sufficient condition for P; we can then show for almost any event C, that C is an inus condition for P ... For ... $C(\bar{C} \vee A)B$ is also minimal sufficient.[30]

For example, let C be the quasi-event that Harry took the righthand turn, let D be the quasi-event that Harry took the lefthand turn and/or Harry failed to take his road-map , let A be the quasi-event that Harry failed to take his road-map, let B be the quasi-event that the signposts had been tampered with and let P be the quasi-event that Harry is late. Suppose that, for some W, there is the PPR,

Pres: AB, F; $[AB] \vee [W] \leftarrow [P]$,

then, presumably A and B cause P.

Now suppose that C occurred, but that C is causally irrelevant to P. Suppose also that it is part of every normal state-of-affairs that Harry either takes the righthand or the lefthand turn. Then D is *equivalent* to $\bar{C} \vee A$, so there is a PPR,

Pres: BCD, F; $[BCD] \vee [W] \leftarrow [P]$,

yet C is not a cause of P.

Reply: The quasi-event BD is not a *regular* part of BCD, for if N is some normal state of affairs $N \times BD$ is not defined.[31] This reply is, roughly speaking, anticipated by Kim; disjunctive quasi-events are excluded. Notice, however, that the quasi-event BD is not

---

30. Kim, J., "Causes and Events: Mackie on Causation", *Journal of Philosophy*, vol58(July 1971), p. 433.

31. See p.43 for the definition of a *regular* part of a quasi-event.

excluded on purely *formal* grounds. It would not affect my reply if D were described as 'Harry did not take a right-road' where the word 'right-road' is used of righthand turns and road-maps. Thus this reply does not make causation language-dependent. (I have already assumed, of course, that causation is relative to some field.)

7.2 Objection Two:

In the case of symmetric overdetermination,[32] one has a PPR,

Pres: C ∨ D, P; [C ∨ D] ∨ [E] ↔ [P],

where C and D overdetermine P. So, on the Regularity Account, C ∨ D should *cause* P. Yet one does not ordinarily speak of disjunctive state-of-affairs as causes. For example, suppose throwing down a cigarette butt and a lightning strike are individually sufficient for the straw to burn and both events occur. Then one would not say that the cause of the straw burning was the fact that the cigarette end was thrown down and/or the lightning struck.

Reply: It would be easy to alter *Reg 1* so that the reductians is false if, in the reductiandum, the proposed cause is disjunctive. For in all three cases, condition (2) could be altered so that there *must* be some X (perhaps A itself) such that AX is the same as C and both A and X are regular parts of C.[33]

One could then argue that C ∨ D cannot be the same as (C ∨ D)X, where C ∨ D and X are regular parts of C ∨ D, since C ∨ D is not a regular part of itself. For if N is a norm N × (C ∨ D) is not, in general, defined. In the example of the straw burning there is no *unique* determinate state-of-affairs obtained by altering the norm to ensure that lightning strikes and/or the cigarette butt is dropped.

---

32. As in Chapter Two, examples of symmetric overdetermination are to be "charitably" interpreted - see the note on p.61.

33. See pp. 154, 155.

However, if one denies that C and D are causes of P, then surely P has some cause and C ∨ D is the most suitable candidate for a cause of P. So *either* one should reject the necessity of the Necessity Thesis *or* one should admit that C ∨ D is the cause of P. If C ∨ D causes P, then it is not the notion of cause that has been used in an unfamiliar way; rather, it is the unfamiliar notion of a disjunctive state-of-affairs which makes the causal sentence seem peculiar.

7.3 Objection Three:

If there is a PPR,

Pres: C, P; [C] ∨ [D] ↔ [P],

there is *also* a PPR,

Pres: C ∨ D, P; [C ∨ D] ↔ [P].

Yet where D does not occur, surely C ∨ D does not cause P. For example, if the lightning-strike caused the straw to burn and no cigarette butt was thrown down, one would not say that the disjunctive state-of-affairs (that the lightning strikes and/or a cigarette butt is thrown down) caused the straw to burn.

Reply: Some rule is required to *select* the suitable PPR out of all those that occur. This selection rule - to be stated in Section Eight - is based on the principle that in the PPR,

Pres: C, P; [C] ∨ [D] ↔ [P],

the "minimal sufficient condition" C should contain as many relevant quasi-events as possible. In other words the relevant circumstances should be considered. In particular, C contains C ∨ D but C ∨ D does not contain C, so the PPR,

Pres: C, P; [C] ∨ [D] ↔ [P]

if preferred to the PPR,

Pres: C ∨ D, P; [C ∨ D] ↔ [P].

7.4 Objection Four: The Problem of Collateral Effects (Part One).

The problem, as stated by Mackie,[34] is that if C causes both A and B and A is earlier than B one has regularities Reg: [CX] ∨ [Y] ↔ A and Reg: [CZ] ∨ [W] ↔ [B] and hence the regularity Reg: [A$\bar{Y}$Z] → [B]. Yet A is not a cause of B. Mackie gives the example in which C is the quasi-event that is it five o'clock, A is the quasi-event that Manchester factory-hooters sound and B is the quasi-event that Londoners stop work. Mackie assumes that one does *not* simply have the regularities Reg: [C] ↔ [A] and Reg: [C] ↔ [B] but that the regularities are of the form Reg: [CX] ∨ [Y] ↔ [A] and Reg: [CZ] ∨ [W] ↔ [B], where on the occasion in question A, B, C, X and Z occur but Y and W do not.

Later (in Section Ten) I discuss the special case in which there is the regularity Reg: [C] ↔ [A]. Notice also that Mackie's argument leads only to the conclusion that A is sufficient-in-the circumstances for B. For one does not have Reg: [A$\bar{Y}$Z] ↔ [B] but only Reg: [A$\bar{Y}$Z] → [B]. Also the absence of a quasi-event need not be mentioned in the statement of regularities unless the absence of a quasi-event is itself a quasi-event. A restatement of the problem is as follows:

If A and B are *collateral* effects of C, one has the PPRs,

Pres: CX, A; [CX] ∨ [Y] ↔ [A]

and Pres: CZ, B; [CZ] ∨ [W] ↔ [B]

If there is a quasi-event ~Y, there is the derived PPR,

Pres: (~Y)AZ ∨ CZ, B; [(~Y)AZ ∨ CZ] ∨ [W] ↔ [B]

If $\bar{Y}$ is *merely* the absence of a quasi-event, there is the derived PPR,

Pres: AZ ∨ CZ, B; [AZ ∨ CZ] ∨ [W] ↔ [B].

In either case, on the account *Reg 1*, there should be a disjunctive

34. Mackie, J.L., *The Cement of the Universe*, pp. 83,84.

cause, (~Y)AZ ∨ CZ or AZ ∨ CZ of P. Yet A and C do not *overdetermine* B. Thus the problem for the regularity theorist is how to distinguish *collateral effect* situations from cases of *symmetric overdetermination*.

Reply: The selection rule of Section Eight shows that one is to prefer the PPR

Pres: CZ, B; [CZ] ∨ [W] ↔ [B]

to the PPR

Pres: (~Y)AZ ∨ CZ, B; [(~Y)AZ ∨ CZ] ∨ [W] ↔ [B]

or Pres: AZ ∨ CZ, B; [AZ ∨ CZ] ∨ [W] ↔ [B],

because CZ contains (~Y)AZ ∨ CZ and AZ ∨ CZ but not vice versa.

7.5 Objection Five: A Problem with Causal Chains

Suppose that C causes A, which, in turn, causes B and that C causes B. Then there are the PPRs

Pres: CX, A; [CX] ∨ [Y] ↔ [A]

Pres: AZ, B, [AZ] ∨ [W] ↔ [B]

and Pres: CXZ, B; [CXZ] ∨ [YZ ∨ W] ↔ [B].

In addition, there is the derived PPR,

Pres: CXZ ∨ AZ, B; [CXZ ∨ AZ] ∨ [YZ ∨ W] ↔ [B].

Yet C and A do not overdetermine B, so CX ∨ A is not a cause of B.

Reply: Notice that this problem is formally similar to the problem of collateral effects. This suggests that a solution in terms of causal priority, such as Mackie provides, would not be required for the problem of collateral effects, since considerations of causal priority would not seem relevant to the problem with causal chains. The solution is to use the selection rule to be stated in Section Eight; CXZ ∨ AZ is part of CXZ but not vice versa.

7.6 Objection Six: The Problems of Overdetermination.

If one accepts the necessity of the Necessity Thesis the reduction *Reg 1* is successful in the case of *symmetric* overdetermination. Cases of *temporal asymmetries* may be handled (as in Chapter Two) by a careful choice of the paraphrase of causal sentences into sentences which explicitly mention quasi-events.

The interesting case is that of *non-temporal asymmetries*. For in such cases A and B overdetermine C and there is the PPR,

Pres: AX ∨ BY, P; [AX ∨ BY] ∨ [W] ↔ [P].

Yet A but not B is said to cause C.

Discussion: In cases of non-temporal asymmetries there are two possible causal chains one of which is broken. Suppose that A causes P (directly or indirectly) but that the causal chain B to D to P is broken at D, that is, D failed to occur, but if D had occurred D would have been sufficient for P. In that case one has the PPR,

Pres: AX, P; [AX] ∨ [DZ ∨ U] ↔ [P],

and the regularity,

Reg: [BT] ∨ [V] ↔ [D].

Also, one may assume that the circumstances Z occur so that if D occurred D would have been necessary and sufficient in the circumstance for P.

Furthermore, since there would be a causal chain B to D to P if D occurred, there is the regularity,

Reg: [AX] ∨ [BTZ] ∨ [VZ] ∨ [U] ↔ [P].

On the one hand, suppose that D fails to occur because T fails to occur. There is the PPR,

Pres: AX, P; [AX] ∨ [BTZ ∨ VZ ∨ U] ↔ [P]

which - as in the discussion of Objection Three - is preferred to the

PPR,

Pres: AX ∨ BTZ, P; [AX ∨ BTZ] ∨ [VZ ∨ U] ↚ [P].

On the other hand, suppose that D fails to occur even though B and T occur and are *almost* sufficient for D. In this anomalous situation there is *not* the PPR,

Pres: AX, P; [AX] ∨ [BTZ ∨ VZ ∨ U] ↚ [P],

because BTZ occurs.

The solution to the problem is to use the selection rule of Section Eight and to prefer the PPR,

Pres: AX, P; [AX] ∨ [DBTZ ∨ VZ ∨ U] ↚ [P]

to the PPR,

Pres: AX ∨ BTZ, P; [AX ∨ BTZ] ∨ [VZ ∨ U] ↚ [P],

AX contains AX ∨ BTZ, but not vice versa.

Section Eight: A Selection Rule

As has been shown in the replies to Objections Three to Six, one has to select a PPR in which the "minimal sufficient condition" that actually occurs contains as many relevant quasi-events as possible. I now make some definitions as a preliminary to stating the selection rule.

Definition: The PPR,

Pres: C, P; [C] ∨ [D] ↚ [P]

is said to *dominate* the PPR,

Pres: A, P; [A] ∨ [B] ↚ [P]

if and only if

(1) D and P are distinct existences;
(2) C is no earlier than A;
(3) C is earlier that P, or, if A and P are simultaneous, C is no later than P.
(4) C contains A.

Definition: A PPR is said to be *complete* if there is no other PPR which dominates it.

Example: Suppose that there is a causal chain A to B to P. Then there are PPRs,

(1) Pres: AX, B; [AX] ∨ [Y] ↞ [B]

(2) Pres: BZ, P; [BZ] ∨ [W] ↞ [P]

(3) Pres: AXZ, P; [AXZ] ∨ [YZ ∨ W] ↞ [P]

(4) Pres: ABXZ, P; [ABXZ] ∨ [YZ ∨ W] ↞ [P].

PPR (4) *dominates* PPR (3), so PPR (3) is *not complete*.

The Selection Rule

The selection rule states that only *complete* PPRs are to be used in the reduction of *causation*. With this proviso, *Reg 1*[35] is converted into *Reg 2*.

Note One: Condition (1) in the definition of *dominance* is required to prevent one arguing that, if A and P are simultaneous, the PPR,

Pres: A(~B),P; [A(~B)] ∨ [BP] ↞ [P],

dominates the PPR,

Pres: A, P: [A] ∨ [B] ↞ [P].

Note Two: Conditions (2) and (3) in the definition of *dominance* are used to restrict the circumstances being considered to quasi-events occurring about the time of the cause and the effect.

Note Three: Completeness is a property of PPRs not of regularities. This complication is unavoidable. For suppose that C*, D* and P* are *qualitatively identical* to C, D and P respectively. Suppose that, there are PPRs,

(1) Pres: C ∨ D, P; [C ∨ D] ∨ [E] ↞ [P]

35. See pp. 154, 155.

and Pres: C* ∨ D*, P*; [C* ∨ D*] ∨ [E] ↔ [P*]

Suppose also that C, D, P, C* and P* but *not* D* occur. Then PPR(1) is complete but PPR(2) is not complete.

Moreover, if the Necessity Thesis is *necessary* C does not cause P (this is a case of overdetermination) but C* does cause P*. So the relevant property is one of the PPRs and not a property of the regularities which in both cases are the same, namely

Reg: [C] ∨ [D] ∨ [E] ↔ [P]

Note Four: The selection rule has been chosen for two reasons. First, it provides a solution to the problems mentioned in Section Seven. Second, it provides a way of incorporating into the regularity those circumstances which are considered relevant; the use of the selection rule in this chapter roughly corresponds to the consideration of the circumstances in Chapter Two.

Section Nine: Further Objections and an Extra Selection Rule

Even with the above selection rule there are some difficulties with the Regularity Account of causation.

9.1 Objection Seven: Accidental Regularities

The selection rule does not enable one to distinguish between discoverable accidental regularities and non-accidental regularities.[36]

9.2 Objection Eight: Unique Causation

It is conceivable that a certain type of bomb (an N-bomb) is only ever exploded once and is detonated using an A-bomb. There are regularities between nuclear events from which the assertion that every explosion of an N-bomb is preceded by an explosion of an A-bomb could be derived. It would seem to be an ad hoc assertion if the

36. See p. 141.

regularity theorist claimed as part of his hypothesis that such examples of unique causation do not occur.

9.3 The Support of a Regularity

In using the selection rule of Section Eight I have already made the account of 'X causes Y' involve several regularities rather than a single regularity, because the *completeness* of a PPR depends on other regularities. I now consider the way in which one regularity *supports* or *counteracts* another. This will enable me to give an account of the distinction between accidental and non-accidental regularities. My aim is not merely to find marks of the distinction between accidental and non-accidental regularities, but also to give an account of the distinction "in the objects".

9.4 Consider the example of unique causation mentioned in Objection Nine. Quasi-event C (the explosion of the A-bomb) causes quasi-event E (the explosion of the N-bomb). I suggest that there is a regularity Reg: [C ∨ W] ← [E] *with only one instance*. (Perhaps this involves a stipulative redefinition of the term 'regularity'.) The regularity Reg: [C ∨ W] ← [E] deserves to be called a regularity because it is supported by the regularity between nuclear events, or by several such regularities. Accordingly, I propose an extra selection rule:

> Only regularities which have a very large number of instances or which are supported by regularities with a very large number of instances are to be treated as non-accidental regularities.

This extra selection rule is based on the principle that the *marks* of the distinction between accidental and non-accidental *generalisations* (discussed in Chapter Two)[37] constitute the only difference

37. See p. 115.

"in the objects" between the corresponding *regularities*.

Does use of this extra selection rule commit one to Scientific Realism? I suggest that it does *not*. The instrumentalist claims that the theoretical entities of Physics are useful fictions, one of the functions of which might be to give an account of which regularities are accidental. However, there is a *slight* peculiarity about the instrumentalist using this extra selection rule. For suppose that there are two rival, equally useful, scientific theories. The first theory entails that when an A-bomb explodes inside an N-bomb the N-bomb explodes. The second theory entails that N-bombs sometimes explode spontaneously yet the explosion of an A-bomb is quite irrelevant. The scientific realist would say that he merely did not *know* whether the explosion of the A-bomb caused the N-bomb to explode. The instrumentalist would be committed to denying that the causal sentence 'The explosion of the A-bomb caused the N-bomb to explode', has a truth-value. I do not consider that this commitment of the instrumentalist is sufficiently peculiar for him to reject the extra selection rule.

Note: Regularities Reg_1, ..., Reg_n *support* a regularity Reg_o if there is some fairly *simple* theory of which Reg_o, Reg_1, ..., Reg_n are consequences and there is no *simpler* theory of which Reg_o but not all of Reg_1, ..., Reg_n are consequences. By a regularity being a *consequence* of the theory, I mean that the regularity is entailed by the laws of Nature (or non-accidental generalisations) proposed by the theory *together with* the correspondence rules of the theory. One account of the *simplicity* of theories, which seems suitable in this context, is Sober's.[38] On Sober's account the less the *extra information* required to answer questions (given the theory) the *simpler* is the

38. Sober, E., *Simplicity*, Chapter One, pp. 1-37.

theory. In this case, the appropriate questions would be of the form: 'Will quasi-event P occur?'

9.5 Next I shall consider *accidental* regularities. In the example of the bombs with red and blue spots where only the bombs with red spots explode,[39] I suggested that the accidental regularity is *outweighed* by a statistical regularity stating the independence of nuclear processes and a wide variety of macroscopic events or conditions. Accordingly, I define a *statistical regularity* to be the occurrence of a large number of quasi-events qualitatively identical to Y, a proportion α of which are preceded by (or simultaneous with) quasi-events qualitatively identical to X. I write this,
Stat Reg: $[X] \leftarrow (\alpha)[Y]$. The statistical regularity is said to be a *basic* statistical regularity if there is no statistical regularity Stat Reg: $[X \vee W] \leftarrow (\beta)[Y]$, which is not too complicated and is such that β is *significantly* greater than α. Similarly, one can consider basic statistical regularities between kinds of quasi-event,
Stat Reg: $K \leftarrow (\alpha)\ J$ where K and J are kinds of quais-events. If Stat Reg: $K \leftarrow (\alpha)\ J$ is a basic statistical regularity between kinds of quasi-event and α is significantly less than 1, then
Stat Reg: $K \leftarrow (\alpha)\ J$ tends to outweigh a regularity Reg: $[X] \leftarrow [Y]$ where X is of kind K and Y of kind J. I suggest that if Reg: $[X] \leftarrow [Y]$ has only *moderate* support (that is, it has only at most moderately many instances and is not supported by regularities with more than a moderate number of instances) but Stat Reg: $K \leftarrow (\alpha)J$ has very many instances and α is significantly less than 1, then the regularity Reg: $[X] \leftarrow [Y]$ should be deemed *accidental*.

The Regularity Account, *Reg 2*, is now modified so that both these

39. See p. 141.

extra selection rules are used. In this way *Reg 3* is obtained.

9.6 An alternative account of the difference between accidental and non-accidental regularities can be provided. One first reduces the regularities into the terms of the current physical theory (involving sub-microscopic events). Presumably all macroscopic regularities then become extremely complicated. Those that are supported by *simple* regularities (involving sub-atomic events) are deemed *non-accidental* But one deems *accidental* those that are not. Notice that, in this case, one is not comparing two observationally equivalent hypothesis for simplicity. The hypothesis in which these accidental regularities are included has more observable consequences and - by accident - has been verified more often than the hypothesis in which the accidental regularities are not included. So in this case the hypothesis which is in other respects a better hypothesis is rejected as being too complicated. Thus one compares the simplicity of the hypothesis with some standard of simplicity. If, for instance, one follows Sober's account of simplicity, one would adapt it by comparing the additional information required to answer questions such as 'Will quasi-event P occur?' with some *standard* of a reasonable amount of additional information.

Section Ten: The Problems of Simultaneous Causation and Collateral Effects (Part Two)

The Regularity Account of causation, *Reg 3* permits cases of *simultaneous causation*. However, as in Chapter Two, in proposed examples of simultaneous causation the account sometimes fails to distinguish cause from effect in cases in which cause and effect are

successfully distinguished pre-theoretically. There is also a special case of the problem of *collateral effects*. Thus there are two further objections to the Regularity Account.

10.1 Objection Nine: Simultaneous Causation

Consider the examples of the *ball on the cushion* and the *hot iron glowing*. In both these examples there are PPRs which, according to *Reg 3*, should result in mutual causation.

Discussion: In Chapter Two (Section Seven) cases of simultaneous causation in which, by *Nec 4*, A and B should cause each other were handled by considering what would happen if an earlier necessary condition for A were removed but if B were *otherwise* made to occur. If A would still occur then B causes A.

The analogue for the Regularity Account would be to suggest the following rule for those cases in which, on *Reg 3*, A and B should cause each other.

Rule One: If there is a PPR,

Pres: C, A; $[C] \vee [D] \leftarrow [A]$

and a regularity,

Reg: $[\bar{BC}] \rightarrow [A]$,

then if, on *Reg 3*, A and B should cause each other, B causes A. For example, if the ball being on the cushion (quasi-event B) causes a depression (quasi-event D), there are PPRs,

Pres: B, D; $[B] \vee [W] \leftarrow [D]$,

and Pres: D, B; $[D] \vee [Z] \leftarrow [B]$,

for suitable W and Z. Why does one say that B causes D and not vice versa? Let C be the quasi-event that, a little earlier, the ball is

just above the cushion. Let E be the quasi-event that, a little earlier, the cushion is not depressed. There are PPRs,

Pres: C, D; [C] ∨ [~E] ∨ [X] ↔ [D]

and Pres: C, B; [C] ∨ [~E] ∨ [Y] ↔ [B].

There is also the regularity,

Reg: $[B\overline{C}] \rightarrow [D]$,

because wherever a ball gets onto a cushion in some other way there is a depression. But there is *not* the regularity

Reg: $[D\overline{C}] \rightarrow [B]$,

because often there are depressions in cushions without balls on them.

10.2 However, Rule One is not satisfactory; consider the *hot iron glowing* example. There are PPRs,

Pres: HI, G; [HI] ∨ X ↔ [G]

and Pres: G, HI; [G] ↔ [HI].

If P is the quasi-event that an H-process has been used, there are the PPRs

Pres: PI, HI; [PI] ↔ [HI]

and Pres: PI, G; [PI] ∨ X ↔ [G].

But there is *neither* the regularity

Reg: $[(\overline{PI})HI] \rightarrow [G[$

nor the regularity

Reg: $[(\overline{PI})G] \rightarrow [HI]$.

Yet HI causes G, not vice versa.

In this case there are no instances of $[(\overline{PI})HI]$, that is $[\overline{PHI}]$; but there are instances of $[(\overline{PI})G]$, namely the cases in which X occurs. (X might be the quasi-event that the material is charcoal and that it is burning.) Rather than permit "regularities with no instances", I consider almost universal generalisations, replacing 'Reg' by 'AUG'.

The difference is merely that an almost universal generalisation is vacuously true if it has no instances. The above discussion of the *hot iron glowing* example now suggests the following rule:

Rule Two: If, according to *Reg 3*, A and B should cause each other but there are PPRs

Pres: C, A; $[C] \vee [D] \leftarrow [A]$

and Pres: C, B; $[C] \vee [E] \leftarrow [B]$

then *If* there is *also*

AUG: $[\overline{C}A] \rightarrow [B]$, *but not*

AUG: $[\overline{C}B] \rightarrow [A]$,

A causes B not vice versa.

10.2 However, Rule Two is also unsatisfactory. For let H be the quasi-event that a substance is heated to 1000°, let S be the quasi-event that the substance is sodium and Let D be the quasi-event that the characteristic yellow light of sodium is emitted. It would be *rash* if - as part of one's metaphysical hypothesis - one denied that Reg: $[SH] \leftrightarrow [D]$ could occur. There are the PPRs.

Pres: SH, D; $[SH] \leftrightarrow [D]$

and Pres: D, SH; $[D] \leftrightarrow [SH]$

let P be the quasi-event that an H-process has been used. Then there are PPRs

Pres: SP, SH; $[SP] \leftrightarrow [SH]$

and Pres: SP, D; $[SP] \leftrightarrow [D]$

but, in this example,

both AUG: $[(\overline{SP})D] \rightarrow [SH]$

and AUG: $[(\overline{SP})SH] \rightarrow [D]$

are *vacuously true*. Yet heating a sodium compound causes it to emit

yellow light, not vice versa.

The significant difference between the two almost universal generalisations is that the latter has greater *support* than the former. For there are two regularities which support AUG: [($\overline{SP}$)D] → [SH], *namely*,

Reg: [D] → [SP]

and Reg: [D] → [SH]

(The first supports the *vacuity* of AUG: [($\overline{SP}$)D] → [SH])

There are *also* two regularities which support AUG: [($\overline{SP}$)SH] → [D], *namely*,

Reg: [SH] → [SP]

and Reg: [SH] → [D].

Now Reg: [D] → [SH] and Reg: [SH] → [D] have the same number of instances *but* Reg: [SH] → [SP] is itself supported by Reg: [H] → [P] which has far more instances than Reg: [D] → [SP], since many other substances are heated.

Note: This argument based on the support of an almost universal generalistation is similar to Gasking's argument in his discussion of the *hot iron glowing* example. There is a *general method* for heating objects and in the *special case* where the object is iron this general method produces a glow.[40]

Accordingly, I propose the following version of the Regularity Account.

Reg. 4

Step One: If, according to *Reg 3*, A causes B but B does not cause A, then A causes B.

---

40. Gasking, D., "Causation and Recipes", *Mind*, vol.64 (Oct 1955), p. 479.

Otherwise

Step Two: A causes B (and not vice versa) if for some C:

(1) There are complete PPRs

Pres: C, A; $[C] \vee [X] \leftarrow (\alpha)[A]$

Pres: C, B; $[C] \vee [Y] \leftarrow (\alpha)[B]$,

where, in most cases, $\alpha = 1$.

and (2i) There is

AUG: $[\overline{C}A] \rightarrow [B]$

but not

AUG: $[\overline{C}B] \rightarrow [A]$

or (2ii) There are

both AUG: $[\overline{C}A] \rightarrow [B]$

and AUG: $[\overline{C}B] \rightarrow [A]$

but the former has greater support than the latter.

Otherwise neither A causes B nor B causes A.

Note One: C need not be necessary in the circumstances for A and B. I suggest that it suffices that C increases the chance of A and B occurring. Hence α (in condition (1)) need not be equal to 1.

Note Two: For a *perfectly symmetrical* case, consider the two yellow lines in the spectrum of sodium compounds. Neither causes the other.

10.4 Objection Ten. Collateral Effects (Part Two).

There are special cases of collateral effects in which one has Reg: $[C] \leftrightarrow [A]$ and Reg: $[CX] \vee [Y] \leftrightarrow [B]$. These in fact, seem only to occur when C and A are simultaneous.

Example: Let C be the quasi-event that the Sun, Moon and Earth are in a certain position relative to each other (which in fact results in a solar eclipse at Paris). Let A be the quasi-event that there is

a solar eclipse at Paris. Let B be the quasi-event that there is a spring-tide (say at Le Havre). Let Y be the quasi-event that Sun, Moon and Earth are in some position, relative to each other, which in fact results in a spring-tide but not in an eclipse. It is assumed that A and Y are described without any mention of solar eclipses or spring-tides. Then there are PPRs

Pres: C, A; $[C] \leftrightarrow [A]$

and Pres: C, B; $[C] \vee [Y] \leftrightarrow [B]$.

Hence there is the PPR

Pres: AC, B; $[AC] \vee [Y] \leftrightarrow [B]$,

yet A does not cause B; rather, C causes both A and B.

Discussion: *Reg 4* requires a slight modification, similar to that used to obtain *Reg 4* from *Reg 3*. The modified account is:

Reg 5: If the conditions for *Reg 4* are satisfied A causes B, *unless* there is some quasi-event C such that:

(1) On *Reg 4*, C should cause both A and B.

(2) There is a regularity Reg: $[C] \leftrightarrow [A]$.

In that case, one compares the generalisations:

(1) AUG: $[(\overline{CX})AX] \rightarrow [B]$

and (2) AUG: $[(\overline{AX})CX] \rightarrow [B]$

Both these generalisations are *vacuously true* and they are supported by:

(3) Reg: $[AX] \rightarrow [B]$

and (4) Reg: $[CX] \rightarrow [B]$, respectively.

If regularity (3) has greater support than regularity (4), there is a *causal chain* C to A to B. Otherwise A and B are *collateral effects* of C.

In the above example, the regularity that a suitable position of

Sun, Moon and Earth is correlated with spring-tides has many more instances than the regularity that a suitable position of Sun, Moon and Earth is correlated with eclipses.

Note: *Reg 5* seems very complicated. However, it is not based merely on an ad hoc reply to various objections, but on the consideration of the producing paradigms in the puzzle cases.

Section Eleven: Conclusions

(1) The Regularity Account (*Reg 5*) satisfies the Reduction-criterion, but not the Analysis-criterion.

(2) Using this account 'X causes Y' is supervenient on:

(i) Y being a distinct existence from the conjunction of X and all earlier quasi-events;

(ii) Various regularities and, perhaps, statistical regularities which are fairly simple and which have large supports.

(3) This version of the Regularity Account involves *several* regularities for each causal sentence, not merely one.

(4) Using the second of my two accounts of the distinction between accidental and non-accidental regularities[41] there is no need to consider, in the regularities, any *kinds* of quasi-event other than *qualitative identity* classes.

(5) Presumably the relation of distinct existence and the simplicity of a regularity are themselves *supervenient* on the quasi-events and the regularity respectively. (It is inconceivable that two possible universes differ only in the relation of distinct existence or only in the simplicity of a regularity; there must be some *other* difference in the quasi-events related

---

41. See p. 169.

or in the regularity.) Hence 'X causes Y' is, I claim, supervenient on X, Y, the quasi-events occurring before X and Y, and the regularities. Thus the Regularity Account provides a vindication of Causal Anti-realism.

(6) The regularity theorist is, I suggest, committed to including in his metaphysical hypothesis the following:

(i) The regularities which (according to the Regularity Account) are not accidental are such that the corresponding generalisations *and* the initial state of the universe (or, if the universe has no beginning, the occurrence of the quasi-events which occur before some time t_o) *jointly entail* that those generalisations have sufficiently many instances (or sufficiently great support) to be considered regularities.[42]

(ii) There is no distinction "in the objects" between non-accidental regularities and in principle undiscoverable accidental regularities.[43]

---

42. This is the regularity reduction of the weaker of the two assumptions stated on p.136 in terms of *fixity*

43. See p. 140.

CHAPTER FOUR

BACKWARDS CAUSATION

In Chapter Two and Chapter Three I ignored the possibility of backwards causation. But the methods used to distinguish cause and effect in cases of simultaneous causation might also, if one wished, provide an account of when backwards causation occurs. For example, consider Mackie's chocolate machine M, which always produces a chocolate bar when a shilling is put in, but which sometimes also produces a free chocolate bar. Someone might argue that the chocolate bar coming out was necessary in the circumstances for the shilling to be put in, and so, if backwards causation is possible, the chocolate bar coming out *causes* the shilling to go in. This is of course absurd. In order to avoid such absurdities one might require that if Y is earlier than X, then Y is causally dependent on X if and only if:

(1) X is necessary in the circumstances for Y;

and (2) For any cause Z of X earlier than Y, if Z does not occur but Y does occur then, in the circumstances, X does not occur.

So, if X is later than Y but X causes Y, Y must not be sufficient for X in the *special circumstances* in which all steps have been taken to prevent X that could have been taken before Y occurs.

However, instead of developing the details of such an account, I shall argue for two theses which jointly show that no proposed case of backwards causation is relevant to the problem of giving an account of causation. The first thesis is that, unless one uses some distinctively *metaphysical* premiss, one cannot show that backwards causation is impossible. The second thesis is that there are no

coherent (hypothetical) examples of backwards causation which could not be rationally interpreted by a causal antirealist without using the notion of backwards causation. (Some of these examples a causal realist *would, I suggest,* interpret as cases of backwards causation.)

In Section One I argue for the first thesis, namely that one cannot show that backwards causation is impossible. In Section Two I briefly expound Mackie's discussion of backwards causation in *The Cement of the Universe* (Chapter Seven, "The Direction of Causation").[1] In Section Three I make some remarks which are relevant to Section Four, where I consider a variety of proposed hypothetical examples of backwards causation and I argue that they bear plausible *forwards causal* interpretations. Throughout this chapter I consider *producing,* not explanatory, causation.

Section One: The Possibility of Backwards Causation

First, I examine three arguments against backwards causation, which are due to Black, Gale, and Swinburne respectively. In each case I do not merely object to the argument as it stands, but I show how the kinds of consideration being discussed illustrate some peculiarity - but not the impossibility - of backwards causation. Then I argue that backwards causation is possible provided the metaphysical theory that there is an irreducible element of causal priority "in the objects" is considered logically compatible with generally accepted truths.

1.1 Some Preliminary Points of Clarification

In discussing whether backwards causation is possible I assume that it is not, in any obvious and straightforward sense, part of the meaning of 'cause' that causes are no later than their effects. Thus I assume that even if the sentence 'No cause is later than its effects'

1. Mackie, J.L., *The Cement of the Universe,* pp. 162-192.

is *analytic*, it is *not trivially analytic* in the way that 'No bachelor is married' is trivially analytic. Hence I consider arguments which purport to reveal *hidden* inconsistencies in the notion of backwards causation. In addition I do not consider the strict logical possibility of backwards causation, but rather the logical *compatibility* of backwards causation with generally accepted truths. Moreover, just as I assume that it is not a trivial analytic truth that no cause is later than its effects, I also assume that it is *not a trivial analytic truth* that the causal "beginning" of a process is identical to the temporal beginning of that process. If there were cases of backwards causation which were not mnemic, a backwards causal process would link cause and effect. Finally, I do not defend the *absurd* thesis that the past can be altered in any but a counterfactual sense of 'alter'. Using the counterfactual sense of 'alter' one says that the past (or the future) is altered by an event C if, but for the event C, the past (or the future) would have been (or would be) different. If a cause C is later than its effect E then C only alters the past in the counterfactual sense, just as if C is earlier than E, C only alters the future in the counterfactual sense. Hence, I submit, backwards causation does not entail altering the past in any but the counterfactual sense of 'alter'.

1.2 Black's Argument[2]

Black presents a hypothetical case of "backwards causation", which he considers to be typical of proposed examples; Houdini, if hypnotized, can always tell how a tossed penny will fall. Here Houdini's answer (event A) might be proposed as an *effect* of the result of the toss (event T) a minute later. To exclude obvious alternative descriptions of the situation in terms of forwards causation Black

2. Black, M., "Why cannot an Effect Precede its Cause?", *Analysis*, vol.16 (Jan 1956), pp. 49-58.

stipulates that A is not a (partial) cause of T, so that the occurrence of A in no way influences the occurrence of T. Black also stipulates that event T has a sufficient cause which is not a sufficient cause of A, so that T and A are not collateral effects of some earlier event X. In addition, he stipulates that the event A is not caused by any *prior* event. This third stipulation is designed to exclude the case in which A is caused by X and T is caused by Y, where X and Y are not causally connected. Some care needs to be taken in interpreting this third stipulation. Perhaps a partial cause of Houdini's giving the right answer might be that he is hypnotized. Black, surely, only wants to deny that there could be a *sufficient* cause of A, earlier than A, which determines A. (If there were such a sufficient cause of A then that cause and T might overdetermine A and so, if the case were deemed symmetrical, T does not cause A. But, presumably, this peculiar case of overdetermination would still be considered a kind of backwards causation for the purpose of this discussion.) Therefore, I accept Black's stipulation as relevant, provided it is a stipulation against there being some earlier event X which is *sufficient* in the circumstances for A.

Black argues that T does not cause A because someone could wait for A to happen and could then prevent T. The obvious answer to Black's argument is to say that if backwards causation occurs, then one *cannot* intervene: it is fixed by some earlier event D that no agent can intervene to prevent T. However this reply is not satisfactory in the case where it is a generally accepted truth that a person *can* intervene in such situations unless there is some kind of mechanical obstruction preventing him. Thus Black's argument would seem to show that backwards causation is incompatible with generally accepted

truths.

Accordingly, I shall consider in greater detail the claim that someone could intervene to prevent T, after A has occurred. There are three senses in which the phrase 'someone could intervene to prevent T' could be understood.

(1) One might merely claim that it was logically possible that some agent (God, for example) intervenes to prevent T, but one might not claim that in fact there is any agent who has the power to intervene. The mere *logical* possibility of preventing T after A occurs does not in any way threaten the claim that C causes A. Compare the case of forwards causation; event C causes a later event E but it is logically possible that, in the circumstances, C did not occur yet E still occurred (spontaneously). Indeed, unless there is this logical possibility, C and E are not distinct existences. However, if one claims that no agent has the power to intervene to prevent T, then one has denied a generally accepted truth about the powers of human beings; unless he is obstructed by some mechanism, a human being has the power to prevent a coin falling to the ground.

(2) Suppose that there is someone who is free to intervene to prevent T in the libertarian sense of 'free', that is, suppose there is no *transeunt* cause of his failure to intervene. In this case, if he intervenes after A has occurred then T does not occur, so T does not cause A. Thus A has *either* occurred spontaneously *or* A has some other sufficient cause (earlier than A). But it is assumed that A has no sufficient cause earlier than A, so if the person intervenes A occurs spontaneously. Thus either A would occur spontaneously anyway, in which case T does not cause A, or the person can intervene after A occurs to ensure that A occurs spontaneously. But the latter

alternative implies that the person has some god-like power of determining laws of nature (and what causes what) and it is generally agreed that human beings do not have such powers. Backwards causation in this case would be incompatible with generally accepted truths.

(3) The sense of 'power' in which it is a generally accepted truth that people do have the power to intervene to prevent T is such that possession of a power to act is *compatible* with the action having a sufficient-in-the-circumstances transeunt cause. Suppose that, in situations in which backward causation occurs, Compatibilism is correct. Thus people have the power to intervene but there is some quasi-event D which is sufficient, in the circumstances, for the absence of intervention. Let C be the quasi-event that no agent intervenes. Suppose also that the compatibilist claims that an agent's actions are determined by *earlier* events, so D is earlier than C. Because intervention could occur any time after A, D must be no later than A. In particular, D is not T itself. Now suppose that D does not occur, then nothing prevents interference, so as in (2), T might not occur. If T caused A but D did not occur, T and hence A might not occur either. So D seems to be a necessary part of a sufficient condition for A. The other necessary part is, presumably, T itself. Consequently, D and T are jointly sufficient for A. For instance, the quasi-event D might be the state of affairs prior to A in all minds, and so D ensures that Houdini is *trying* to guess the future and is in a state in which he *can* know what will happen. D also ensures that no agent intervenes. Thus there are the following causal connections:

(i) T is necessary for A;

(ii) D and T are jointly sufficient for A (perhaps also D is necessary for A);

(iii) D is sufficient for C;

(iv) C is necessary for T.

To show that such a system of causal connections is possible, I consider the following case of forwards causation:

D* is the kitchen being on fire;

T* is the presence of gas in the house;

A* is the explosion of the house;

C* is the bursting of the gas cylinder in the kitchen.

It is then plausible that D* and T* are individually necessary and jointly sufficient for A*, that D* is sufficient for C*, and that C* is necessary for T*. Nonetheless it is plausible that T* is a *cause* of A*. It follows that Black's argument merely shows that for a case of backwards causation to be consistent with generally accepted truths, some form of Compatibilism should hold in the backwards causal situations.

Note: If an agent performs an action A and A causes an earlier event C, then A cannot be free in the libertarian sense. For if A were free the agent could freely decide, when C has already occurred, not to perform A. (In a sense, the agent can intervene after C to prevent the proposed cause of C, namely A). I call this result *The Corollary to Black's Argument.*

1.3 An Argument of Gale's[3]

Gale argues that if an event L is later than an event E, and if L caused E then it would be *logically impossible* for an agent to have a trace of E (say a memory) when the agent brings L about. Yet it is *logically possible* for an agent to have a trace of E. Gale gives the example in which Jones on learning that Smith, sometime earlier, was in a dangerous situation, "retro-warns" Smith (event L). Smith, in the dangerous situation, has an experience (event E) in which he seems to hear the words which Jones subsequently utters. Suppose L causes E.

3. Gale, R.M., "Why a Cause Cannot be Later than its Effect", *Review of Metaphysics*, pp. 230-231.

Gale claims that since E is earlier than L it is logically possible that Jones has a trace of E and hence that Jones knows that E occurs. But if Jones *intentionally* does L in order to bring E about, it is impossible for Jones to know that E occurred.

Gale's argument (as he himself admitted in a letter to Brier[4]) is based on the confusion between 'p logically imples q' and 'p implies necessarily q'. For the sentence

> If Jones intentionally does L in order to bring E about, it is impossible for Jones to know that E occured,

must be interpreted as

> 'Jones intentionally does L in order to bring E about' logically implies 'Jones does not know that E occurred',

and not interpreted as

> 'Jones intentionally does L in order to bring E about' implies the logical necessity of 'Jones does not know that E occurred'.

An additional defect in Gale's argument is that a person can sometimes know the occurrence of the intentional effect of his action. For example, if I go to the window with the intention of opening it, then in a suitably weak sense of 'know', I know that I shall open the window. (Presumably Jones' knowledge is knowledge only in some weak sense.) This defect is easily remedied; a person who knows the occurrence of E and whose knowledge is due to a *trace* of E (say a memory) cannot *intentionally* bring about E. And it is precisely knowledge due to a trace which Gale considers.

I now present an argument which is valid but the conclusion of which is not as strong as Gale's. Suppose that E occurs and let L*ϕ be an abbreviation for 'ϕ is true and the truth of ϕ is fixed at

4. Brier, B., *Precognition and the Philosophy of Science*, p.3.

time t_1'.[5]

Let e be: 'E occurs at time t_1

Let p be: 'At time t_2 Jones intentionally does L to bring about E at the time t_1'

Let q be: 'At time t_2 Jones knows that E occurs, as a result of a trace of E'.

Consider the following premises:

(1) $L^*(e \supset p)$

(2) It is *logically* necessary that $q \supset e$.

(3) $L^*(q \supset \sim p)$

Premiss (1) is based on the assumption that the only way in which E can occur, given the history up to E, is by Jones retro-warning Smith and on the assumption that Jones could not retro-warn Smith *unintentionally*. Although $L^*(e \supset p)$ is not a logical consequence of a backwards causal-situation in which Jones retro-warns Smith, it would be peculiar not to admit the possibility of premiss(1) and yet to accept the possibility of backwards causation. Premiss (2) is obvious; knowledge that E occurs *entails* that E occurs. Premiss (3) is the premiss that it is impossible both to know that E occurs *as a result of a trace* and intentionally to bring E about. This impossibility might be slightly weaker than logical impossibility, but it is surely strong enough to permit the premiss $L^*\sim(p\&q)$, which is equivalent to Premiss (3).

---

5. It is plausible that L* has the properties of necessity in Fey's system T. It is also plausible that $F^*F^*\phi$ is true, for all ϕ, where F* is the non-contingency operator defined by: $F^*\phi =_{Df} L^*\phi \vee L^*\sim\phi$. It follows that L* has the properties of necessity in Lewis' system S5. (See Montgomery, H.A., and Routley, F.R., "Contingency and Non-Contingency bases in Normal Modal Logics", *Logique et Analyse*, vol.9 (1966), pp. 318-328.

From (2), one has (4), $L^*(q \supset e)$

From (1) and (4), one has (5), $L^*(q \supset p)$

From (3) and (5), one has $L^*(\sim q)$

It follows that it is fixed *at the time of the occurrence of E* that Jones will not know that E occurred when he retro-warns Smith. (At least, any knowledge which Jones has is due to Jones' intention to retro-warn Smith, rather than to a trace of E.) This is indeed a somewhat peculiar result but it does not establish the impossibility of backwards causation. For there could be some event D earlier than E, such that D and L are individually necessary and jointly sufficient for E. D might also prevent Jones knowing that E occurs. For instance, D might be the state of some "Psi field"; only if the Psi field is in a certain state can retro-warning occur, and if the Psi field is in this state, certain memory traces are erased.

1.4 An Argument of Swinburne's[6]

Swinburne's argument against the possibility of backwards causation is based on two premises.

Premiss One: If a statement is to be about a certain instant then it must be logically possible for an observer, who is correctly positioned, either to verify or to falsify the statement more conclusively at that instant than at other instants.

Premiss Two: An action which affects the past produces changes in the evidence about the past. Thus Swinburne assumes that if an action A at time t_2 causes an event B at time t_1, then there is evidence at t_1 that B has not occurred, but at t_3 (after t_2) there is evidence that

---

6. Swinburne, R., *Space and Time*, pp. 157-170.

B has occurred.

Swinburne now argues[7] on the one hand, that if the evidence at t_1 is *better* than the evidence at t_3, then by Premiss Two there are no grounds for believing that B occurred. But, on the other hand, if the evidence at t_3 is better than the evidence at t_1, then

> <u>either</u> the evidence at t_1 could not be improved (so as to be better than that at t_3), therefore by Premiss One the statement that B occurred is *not* about instant t_1;
>
> <u>or</u> the evidence at t_1 was not typical and more evidence would have resulted in a different conclusion, namely that B *had* occurred. In this case, by Premiss Two, B occurred but action A did not *cause* B.

Therefore, Swinburne argues, there could be no grounds for believing that A at t_2 *caused* B at t_1. He then concludes:

> Since no evidence could ever - for reasons of logic - support a claim that a man affected the past it makes no sense to suppose that he could.[8]

Even if one does not accept the positivist-inspired principle Swinburne is using, the impossibility of there being grounds for believing in backwards causation would be an important result. Finally, Swinburne argues that if it is not possible, even in principle, for a person to affect the past, then there could be no cases of backwards causation in which an *event* affects the past. Swinburne here uses the premiss that if event C causes E a human being could *in principle* use C to produce E.

I now discuss the following two objections to Swinburne's argument.

---

7. Swinburne, R., *op.cit.*, p.164.
8. Swinburne, R., *op.cit.*, p.167.

Objection One: When Swinburne discusses the case in which the evidence at t_3 is better than the evidence at t_1, Swinburne introduces a *false dichotomy;* he assumes that *either* the evidence at t_1 is *not typical* (and that more evidence would lead to a different conclusion) *or* the evidence at t_1 is the *best possible evidence* at t_1. Thus Swinburne ignores the case in which the evidence at t_1 is not the best possible but better evidence at t_1 would lead to the same conclusion.

Discussion: Swinburne's argument could be modified using an extra premiss:

Premiss Three: If an action A affects the past then the evidence available after the action A differs from the best possible evidence before the action A.

By Premiss Three, if the actual evidence at t_1 was that B did not occur but better evidence would have shown that B occurred, then action A did not cause event B. (Action A is interpreted as affecting the condition or appreciation of the evidence.)

Objection Two (due to Brier[9]): Swinburne assumes (Premiss Two[10]) that without a change in evidence there could be no grounds for supposing that an action affects the past. While he might perhaps be correct, surely he should argue for this point, which is not an obvious one. Brier, for instance, suggests that a strong enough correlation between actions like A and events like B would provide grounds for supposing that A caused B.

Discussion: A variant of Premiss Two might be supported by the Foreknowledge Principle, namely:

If person P could know at time t_1 that person Q will act at time t_2 (later than t_1)

---

9. Brier, *op.cit.*, p.55.
10. Also, see Swinburne, R., *op.cit.*, p.164.

in order to bring about an event B, then
Q is not free (in the libertarian sense)
to refrain from acting.

This principle can be given more plausibility if the proviso is made that P's knowledge that Q will bring B about is based on evidence available at time t_1 that B will occur, rather than based *either* on a knowledge of Q's intention *or* on a faculty of precognition. The Foreknowledge Principle can be used to support Premiss Two*, namely:

If free action A at time t_2 causes an event
B at time t_1, then there is no evidence available
at t_1 that B has occurred which would be sufficient
to support a *knowledge* claim that B occurred.

But, if the phrase 'best possible' in Premiss Three is interpreted as 'logically best possible', the Foreknowledge Principle could not be used to support Premiss Three. One might try to find some other sense of 'possible' for which both Premises One and Three would be plausible. Consider the following senses:

(1) Logically compatible with all past events;
(2) Logically compatible with laws of nature;
(3) Logically compatible with laws of Nature and all past events.

Using either (1) or (2), Premiss One retains its plausibility, but it is hard to see how Premiss Three could be derived from the Foreknowledge Principle. Using (3), Premiss Three might perhaps be derivable from the Foreknowledge Principle together with other plausible premises. But Premiss (1) is now implausible, for the laws of nature might be deterministic and so only that which actually occurs would be possible in the required sense.

Suppose one were to accept the above modification of Swinburne's

argument. Then one would have shown only that one could never, even in principle, know that someone had *freely* affected the past, in a libertarian sense of 'free'. For, surely, it is possible that person P could in principle at time t_1 know what person Q will freely do at a later time t_2 in any sense of 'freely' which is compatible with Q's free actions being determined by events occurring before t_1.

Thus Swinburne would have established the corollary to Black's argument, namely that in a libertarian sense of 'free' *no agent can freely affect the past.* Swinburne then argues that if C causes E an agent should be able in principle, to use C to bring E about. This argument is incorrect if the action of the agent is required to be free in a libertarian sense. For the paradigms of causation are cases in which an agent acts to produce C and hence E; if such actions must be free in the libertarian sense, then the compatibilist is committed to the extremely peculiar - though perhaps consistent - claim that all the apparent paradigms of causation are *illusory as paradigms* although they might be *like* the *purely imaginary paradigms*. Rather, it seems that the compatibilist would assert that the paradigms of causation are cases in which the action is free only in the compatibilist sense.

Therefore, although Premiss Two* exhibits a *peculiarity* of backwards causation, Swinburne's argument against the possibility of backwards causation fails.

1.5 The Possibility of Backwards Causation

I now sketch an account of the ordinary notion of time which, if it is correct, can be used to argue that backwards causation is *logically possible* but that it is logically necessary that cases of

backwards causation are rare.

I begin by considering three proto-temporal orderings, that is time-like orderings of events.

(1) A person is aware that some perceptual experiences are later than other perceptual experiences. Again, a person has what at least seem to be memories; the remembered experience is earlier than the conscious memory of that experience. It is perhaps *contingent* but presumably in all normal cases *true* that, for any given person, these orderings of perceptual experience and remembered experience cohere into a single proto-temporal ordering. I call this the ordering of *phenomenological time*.

(2) If A causes B then A is *causally prior* to B. A causal proto-temporal ordering is an ordering such that if A is causally prior to B, A is no later than B.

(3) Implicit in the theories of Physics is some account of the geometry of space-time. Typically there are one or two obvious proto-temporal orderings associated with this account of the geometry of space-time. For example, in the theory of Special Relativity there are two possible proto-temporal orderings (one is the reverse of the other) both of which are *partial* for events which are separated in a space-like manner cannot be temporally ordered (except with respect to some particular frame of reference).

Now I suggest that the ordinary notion of time presupposes a great deal of coherence between the phenomenological proto-temporal orderings and some causal proto-temporal orderings, and perhaps a great deal of coherence between both these and some ordering associated

with the geometry of space-time. My only argument in support of this suggestion is that it provides a (transcendental) explanation for the intuition that there is a *logical peculiarity* about backwards causation. I suggest that this logical peculiarity is due to the logical necessity of a great deal of coherence between the phenomenological orderings and some causal ordering if there is to be any temporal ordering at all.

However, if one is a realist about causal priority, one is committed to the *contingency* of this coherence between the causal and the other proto-temporal orderings. Now consider the situation in which there is, on one occasion only, a discrepancy between the causal and the other proto-temporal orderings. For example, suppose that there are many observers of events E and L and, in each case, the experiences of the observers are such that E seems to be before L. Yet L is causally prior to E. Furthermore, suppose that event A is causally prior to event B, that A is spatio-temporally near to E (perhaps they are contiguous) and that B is spatio-temporally near to L. Since this is the only case in which there is a discrepancy between the proto-temporal orderings I suggest that one would not say that the ordinary temporal ordering had become *incoherent*. Furthermore, it would seem most peculiar if one claimed that E is later than L, for in that case A is earlier than B, E is later than L, yet A is near E and B is near L with respect to the geometry of space-time as described in the physical theory. Thus one would interpret this situation as one of backwards causation.

Hence if my account of the ordinary notion of time is correct, the realist about causal priority is committed to the possibility of backwards causation. But since realism about causal priority is logically compatible with generally accepted truths, it follows that

backwards causation is itself logically compatible with generally accepted truths.

Note: In this argument I assume that the ordinary temporal ordering is based on the various proto-temporal orderings which I mention. If one holds that the ordinary notion of time is irreducible, a similar argument to the above holds. For there is no reason why an irreducible causal priority "in the objects" should agree with this irreducible temporal ordering.

Section Two: Mackie's Account of Causal Priority

2.1 Mackie begins by presenting three arguments in support of a *conceptual distinction* between causal and temporal priority.

Argument One[11]: "We seem ready to accept causes which are simultaneous with their effects."

If one assumes, as I do, that there are coherent examples of simultaneous causation, Argument One shows that causes need not be earlier than their effects. In this chapter I am concerned with the possibility of backwards causation. Clearly, if one has an adequate account of simultaneous causation without involving an irreducible element of causal priority "in the objects", then Argument One has no relevance to the possibility of backwards causation.

Argument Two: Mackie argues that backwards causation is conceivable and promises to describe "in some detail conceivable experimental results which, if they were obtained, would be coherently interpreted as evidence for backwards causation."[12]

This argument depends on the success of the promised hypothetical

11. Mackie, J.L., *The Cement of the Universe*, p.161.
12. *Ibid.*, p. 161.

example of backwards causation.

Argument Three: Mackie argues[13] that temporal and causal priority have different logical structures. Here Mackie seems to be arguing from the assumption that the temporal ordering of events is a *total* ordering, that is, any two events are temporally comparable. Mackie also seems to assume that the temporal ordering is compatible with a relation of betweenness derived from the topology of the line. From these assumptions it follows that "once a time-direction has been given to any pair of events, it has been given to the system as a whole".[14] The ordering of causal priority is not a total ordering but merely a partial ordering, that is there can be events A and B which are not comparable with respect to the ordering of causal priority. Mackie concludes that there is a conceptual distinction between temporal and causal priority. Mackie's argument fails because, I suggest, the two assumptions which he makes are false if the temporal ordering is that associated with the geometry of space-time. For, according to the theory of Special Relativity, if A and B are separated in a space-like manner, there are many frames of reference with respect to which A is earlier than B, and there will be some event C such that with respect to some of these frames C is later than B but with respect to others C is earlier than B. Mackie's argument is reminiscent of that which I propose in Section 1.5; both arguments are based on the logical possibility of discrepancies between different *time-like* orderings. So it is significant that Mackie says

> An adherent of the latter view [that causal priority is temporal priority in disguise] would say that when it is settled that D is temporally prior to E ... it follows at once that if they

13. Mackie, J.L., *The Cement of the Universe*, pp. 162, 163.

14. *Ibid.*, p.162.

> are causally related, D is the cause and E the effect. But what I am maintaining is that this *seems not to be how we ordinarily see this matter.*[15]

I maintain that how "we ordinarily see this matter" involves an incorrect metaphysical theory which is nonetheless not logically incompatible with generally accepted truths.

Argument Four: Mackie argues in favour of the conceptual distinction between causal and temporal priority on the grounds that "we regard causes as explaining their effects in a way in which effects do not explain their causes, and the mere fact of temporal priority would not account for this".[16]

This argument raises two important issues which will be discussed in Section Three. The first is that the causal anti-realist should give some account of why causes are ordinarily considered to explain their effects, but not vice versa. Otherwise the causal anti-realist's account is suspect (See Section 3.5). The second issue is more directly relevant to the discussion of backwards causation; are there situations in which quasi-events E_1 .... E_n precede L, and of which the only plausible explanation is that E_1 ... E_n are collateral effects of L? This second issue is closely related to the problem of finding the "time-reversed image" of the (forwards-causal) collateral effects situation. Such a "time-reversed image" is, presumably, a forwards-causal description of the situation in which E_1 ... E_n precede L but in which one is tempted to say that E_1 ... E_n are the *backward-causal collateral effects* of L. Mackie argues that the description of the situation as one of causal overdetermination does not explain the coincidence of E_1 ... E_n, but as he himself points out this is not in any case an adequate description of the "time-reversed

15. Mackie, J.L., *The Cement of the Universe*, p.163.
16. *Ibid.*, p.169

image". For example, if there are several precognisers and events E_1, ..., E_n are their precognitions of the same event, L, then L *seems* to be sufficient for E_1, ..., E_n in a way which would not be so if E_1,..., E_n *overdetermined* L. It *seems* that the relevant difference is that (in the precognition example) in the circumstances, if any events D_1, ..., D_n, which are earlier than E_1, ..., E_n, were prevented yet L were ensured, then E_1, ..., E_n would still occur; while in the case of overdetermination E_1, ..., E_n would not occur. Here the circumstances include the quasi-events that the precognisers are trying to precognise and have the ability to precognise. I provide a suitable forwards causal description of this situation in Section 3.6. Thus, assuming that these two interesting issues are treated satisfactorily, not one of Mackie's arguments shows that one cannot incorporate into an analysis of causation the condition that the cause is not later than the effect.

2.2 Having argued that one cannot in a straightforward way identify causal and temporal priority, Mackie considers three accounts of causal priority.

(1) Russell's Account[17]: Russell's suggestion is that if there is a law that entails that an event of kind E occurs if and only if either an event of kind A or an event of kind B or ... occurs, then if an A and an E occur the A causes the E.

Mackie proposes the following counter-example to Russell's suggestion; the indeterministic chocolate machine L should, on Russell's account, involve backwards causation (this example is clearer if the machine says "thank you" if a shilling is put in but no bar comes out ...). However controversial Mackie's chocolate machine might be as an example of forwards causation, it is certainly not an

17. Mackie, J.L., *The Cement of the Universe*, p.166.

example of backwards causation!

(2) The Effectiveness Account: In my Introduction I claimed that the paradigms of producing causation are cases in which a person produces (or prevents) one event in order to produce (or prevent) another. The Effectiveness Acccunt of causal priority is as follows:

> Given X and Y are causally related, X is causally prior to Y if one can produce (or prevent) Y by producing (or preventing) X, but not vice versa.

Mackie has two objections to the Effectiveness Account both of which I have considered in the Introduction. His first objection is that if one considers the hypothetical causal situation in which the approach of another star towards the Sun caused the formation of the planets, then it is clear that such cases of causation are ones in which no human being could produce the cause. But if one claims that in principle an agent could produce planets by making a star approach the Sun, then one's grounds for making this assertion are not based on any experience of making planets. Here it seems one is asserting that in principle an agent could bring about X in order to bring about Y (but not vice versa) *either* precisely on the grounds that X is causally prior to Y, *or* because the situation has some features in common with the producing paradigms of causation. In the former case the effectiveness account is circular, in the latter case it must depend on some other account of causal priority (such as the account which I have given of causation in Chapters Two and Three). So in neither case is the effectiveness account by itself an adequate account of causal priority.

Mackie's second objection is based on Chisholm's example in which

a person raises his arm in order to make certain earlier neural events occur. Mackie claims that on the Effectiveness Account this should be considered a case of backwards causation, which is counter-intuitive. (Although von Wright considers that as part of a suitable "closed system" this would be a case of backwards causation.[18]

(3) Popper's Account[19]:

Mackie also discusses an account of causal priority according to which X is causally prior to Y if X is "prior with respect to the dispersal of order". The paradigm of order-dispersion is the example in which waves spread out from the centre of a pool. This process of expanding waves is explicable if it is the effect of a single event - say a stone being dropped into the centre of the pool. But the reverse process, though physically possible, would involve there being many generators of waves. Popper proposes a Dispersal of Order Account of the direction of time. But, as Mackie points out, it is logically possible that there is a process in which the event which is *prior* with respect to dispersal of order of order is *later* in time; for instance, suppose the waves move inwards from the outside of the pool, without any previous outward movement and without there being coherent generators of waves. Thus dispersal of order seems to be related to causal, rather than temporal priority.

Mackie suggests that priority with respect to dispersal or order is *not itself causal priority* but rather constitutes *evidence for causal priority*. That priority with respect to dispersal of order is not the same as causal priority is, presumably, shown by the logical possibility of many random processes at the edge of the pool

---

18. von Wright, G.H., *Explanation and Understanding*, pp. 74-81.
19. Mackie, J.L., *The Cement of the Universe*, p.183.

happening to result in a coherent pattern of inward-moving waves. That priority with respect to the dispersal of order is *evidence* for causal priority is based on the claim that the situation can only be explained as collateral effects of a common cause; for Mackie says:

> An example of dispersed order, then, such as the outer circular wave, is a set either of collateral effects or of joint causes of the central events. ... Since causes explain effects, and not vice versa, the central event will explain the dispersed order only it it is its cause ... [20]

I argue in Section 3.6 that such apparent cases of collateral effects of a later cause can be interpreted in forwards causal terms.

Thus Mackie has argued that none of the three accounts of causal priority considered (Russell's account, the Effectiveness Account and Popper's account) is satisfactory.

2.3 A Proposed Example of Backwards Causation

Mackie considers a hypothetical precognition experiment based on an example of Scriven's.[21] Mackie does not base his example on the assimilation of precognition to familiar processes of visual perception, and so when he talks about precognition one could substitute (as he suggests) the appearance of a pattern in iron filing on a glass sheet. A pattern is displayed on Tuesday (event B) and one (or more) precognnisers correctly draw that pattern on Monday (or the pattern appears on the sheet of iron filings on Monday). The agreement is too good to be a coincidence. Mackie considers three possibilities.

(1) A causes B;

(2) B causes A;

(3) A and B are collateral effects of

20. Mackie, J.L., *The Cement of the Universe*, p.185.

21. Scriven, M., "Randomness and the Causal Order", *Analysis*, vol.17 (1956), pp. 5-9.

some third event C earlier than A.

Provided that the pattern is selected at random - in the sense of having no cause - the first and the third of these possibilities can be rejected. So it would seem that B causes A. However, as Mackie points out, the notion of a random event itself depends on the notion of a cause. Thus he says "the item which just happens, which pops up from nowhere, is simply something to which nothing is causally prior".[22] This remark is especially relevant to Mackie's search for an appropriate analysis of causal priority. However, it also shows that one is begging the question if one assumes that, say, the radio-active decay responsible for the selection of a pattern is random. Here one might consider a cruder example: the Geiger-counter which registers radiation one minute *before* the radium atom decays - call this the *precognitive Geiger-counter*. Perhaps the clicking of the precognitive Geiger-counter could itself be a random process and cause the radium atom to decay (see Section 3.3).

Mackie then considers Black's argument discussed in Section 1.2. He suggests that Black's argument shows not merely, as I claim, that it is fixed (by an earlier event D) at the time of event A that no one intervenes to prevent B, but also that it is fixed that B occurs. So Mackie excludes the case in which some free choice or random device is used to produce B. But D might, I claim, be sufficient for the non-intervention, yet D and the random event B might be *jointly sufficient* and *individually necessary* for A. Thus it is *not B*, but *rather the randomness of B*, that must be fixed. Indeed, in the case of the precognitive Geiger-counter there is no normal way in which anyone *could* intervene to prevent the radium atom decaying. One might claim that the randomness has *always* been fixed.

---

22. Mackie, J.L., *The Cement of the Universe*, p.177.

Mackie suggests that for an example of backwards causation one requires that A be fixed after B is fixed (but B occurs later than A). Perhaps the random device is used to generate random numbers at time t_0, then it is used, at time t_1, to select the precognisers from the population. The precognisers make their drawings at t_2 and the pattern is selected using the previously generated random numbers and is displayed at time t_3. In this case, the argument for interpreting the situation as one of backwards causation is that the alternative forwards causal explanation is excessively complicated. For it seems to involve a clairvoyant knowledge of the randomly generated numbers and a clairvoyant knowledge of the procedure used to convert the numbers into a pattern. I discuss whether or not this is the only alternative in Section 3.4.

Finally, Mackie suggests that the precognition experiment would be more decisive if there were several precognisers who all make the correct drawing. This argument seems also to depend on the rejection of the alternative interpretation of the situation as one of forward-causal collateral effects.

So Mackie's example raises three interesting issues. The first is whether one should assimilate precognition to, say, visual perception, and treat a case of precognition as one in which the knowledge is the *effect* of the known state of affairs. (I discuss this issue in Section 3.2.) The second issue is whether, where an apparently random device is used, it is reasonable to interpret the experiment as a situation in which the random device is affected by, say, the precogniser. The third issue is whether or not one can interpret the situation in which a precogniser makes a drawing which has *already* been selected but has not yet been displayed, without treating the

precogniser as a clairvoyant.

2.4 Mackie's Account of Causal Priority

I now briefly criticise Mackie's account of causal priority in terms of fixity, which is as follows:

> Suppose that ... X is seen as necessary (and sufficient) in the circumstances for Y ... Then despite this, X was not causally prior to Y if there was a time at which Y was fixed but X was unfixed. If on the other hand X was fixed at a time when Y was unfixed, then X was causally prior to Y. Again if X was not fixed until it occurred, then even if Y [occuring later than X] also was fixed as soon as X occurred, ... X was causally prior to Y. And further, if there is some line or chain of causation ... linking X and Y and some other event Z so that X was between Y and Z, and if Z was not fixed until it occurred, then X was causally prior to Y. 23

I object to Mackie's account of causal priority because it is *ineffective*. For I argue that his account of causal priority is of no use in handling the problems of collateral effects and simultaneous causation which it is partly designed to handle. But if one provides, as I have done in Chapter Two, an analysis which handles these problems without using causal priority, there seems to be no need to use that notion in the analysis of causation.

Consider first cases of collateral effects. None of the examples considered in the discussion of the Necessity Thesis involved random processes and so they are compatible with Total Determinism(that is, the theory that every event has always been fixed). But as Mackie points out, in the case of Total Determinism there is no contrast between fixity and unfixity and so there are no considerations of causal priority as analysed by Mackie.

Again, even if one ignores the possibility of Determinism, there

---

23. Mackie, J.L., *The Cement of the Universe*, p.190.

will still be many hypothetical cases in which X and Y are simultaneous and are *simultaneously fixed* at a time earlier than X or Y. On Mackie's account, to decide whether X causes Y or Y causes X, one has to find an earlier event Z, not fixed until it occurs, such that the causal chain is Z to X to Y, rather than Z to Y to X. But unless one can decide whether X causes Y or Y causes X, one cannot decide *whether* the causal chain is Z to X to Y *or* Z to Y to X.

For example, consider the ball on the cushion. It is supposed that a ball sometimes simply stops above the cushion, or sometimes drops onto the cushion . Let X be the event that the ball is on the cushion, let Y be the event that the cushion is depressed. Let Z be the event that the ball did not stop above the cushion. Then, in this example, Z is only fixed when it occurs; when Z is fixed so also are X and Y. One does not know that the causal chain is Z to X to Y, *rather than* Z to Y to X, until one knows *whether* X causes Y *or* Y causes X. So I conclude that Mackie's account of causal priority is ineffectual.

Note: Presumably, since Mackie is giving an account of causal priority in terms of *fixity*, fixity must itself not involve the notion of causal priority. I suggest that the time at which an event Y is fixed is the time of occurrence of the earliest quasi-event X such that it is nomically necessary that if X occurs Y also occurs.

Section Three: Some Remarks Relevant to the Discussion of Backwards Causation

In this section I gather together some remarks which will be relevant to the examination of proposed cases of backwards causation in the next section.

3.1 On Requiring a Simple Causal Explanation

Suppose that a hypothetical example, say of precognition, is explained in terms of a hypothesis about the causal relations between events, and the simplest such explanation involves backwards causation. Is that evidence for backwards causation? The *causal realist* might treat this as evidence, for by abduction the simplest account of causes "in the objects" would be the one most likely to be correct. But if one is an anti-realist about causal priority and if one can give a *simple alternative* explanation which *does not involve the notion of a cause* but involves only, say, the notion of a law of Nature or of a regularity, then one is *not committed* to backwards causation. In that case, if a causal description of the situation is *demanded*, it would not be irrational to give a *very complicated* description involving forwards causation, provided one has a simple account of what is happening in terms of laws of Nature.

3.2 Precognition without Backwards Causation

If one were to assimilate precognition to visual perception one would claim that the known event causes the knowledge. But I claim that the following conditions are sufficient for a person P to precognise an event E:

(1) E occurs and P believes that E occurs;

(2) The event E is later than P's belief that E occurs.

and (3) P's belief that E occurs is nomically sufficient in the circumstances for the event E to occur. In other words, if 'Nϕ' is an abbreviation for 'ϕ is nomically necessary', if c is a sentence describing the circumstances (up to the time of P's belief), if b is the sentence 'P believes that E occurs' and if e is the sentence 'The event E occurs',

then

$$N((c\&b) \supset (c\&e)).$$

First notice that the question of the believer's *justification* of the belief is not usually considered when discussing precognition. If someone with psychic powers has a vision in which he seems to see the future and in fact, unknown to the psychic, it is a law of Nature that visions of this kind are always correct, then one would say that there was precognition even though the psychic might not be justified in believing that the vision is veridical.

Now consider the case in which an event of kind C invariably causes events of kinds A and B, and events of kind A are only caused by events of kind C. For example, heating a sodium compound produces light of a characteristic yellow colour which has two different wavelengths (the Sodium D lines). If I know the light contains one of the wavelengths, then I know that there has been a sodium compound heated (or otherwise excited) and so I know the light also contains the other wavelength. Yet in this case the known events and the knowledge are collateral effects of heating the sodium. Perhaps in some strict sense the "knowledge" that the unobscured wavelength has occurred is not genuine knowledge. But I claim that precognition would not be required to be genuine knowledge in such a strict sense of 'knowledge'.

Hence I conclude that conditions (1), (2) and (3) are sufficient for P to precognise it and I conclude that one requires *neither* that P's belief be justified *nor* condition (3*) namely:

(3*) Event E is *causally* necessary in the circumstances for P's belief that E occurs.

Note: The difference between conditions (3) and (3*) is most clearly seen in a case of simultaneous causation. Suppose that X and Y are

simultaneous, that X is necessary in the circumstances for Y and that X causes Y. Then Y does not cause X, yet Y is nomically sufficient in the circumstances for X.

3.3 On Affecting "Random" Devices

Suppose that a precogniser correctly draws a pattern which is selected a day later by a device which is usually considered to be random, such as a device which measures the time between the clicks of a Geiger-counter in the presence of a small quantity of a radioactive substance. Assume that the precogniser is always successful and that the success is not an accident. I claim that the situation can be adequately described in terms of laws of Nature, without reference to causation, as follows:

> Given, as the circumstances, the state of mind of the precogniser, P, at the time of the experiment, P's drawing X is nomically necessary and sufficient for the later selection of the pattern X.

Thus I assume that, given P's peculiar mental state, he would not draw the wrong pattern and if no pattern were displayed he would not draw a pattern at all.

Although it is not necessary to do so, one could also describe the situation as one of forwards causation, P's drawing X *causes* the pattern X to be displayed. I now consider two objections to this description of the situation in forwards causal terms.

Objection One: It might be objected that precisely because the device is random the selection of the pattern cannot be caused by the precogniser's drawing.

Reply: *Either* one means by 'random' 'not having a transeunt cause', *or* one means by 'random' 'having an equal propensity to select any pattern', *or* one means by 'random' 'statistically random'. In the first two cases, one should distinguish between an *essentially* random process and a *materially* random process. If it is part of the nature of the device that it must be random then it is said to be *essentially* random. But if the device is random only because it happens to lack a cause then it is *materially* random. If two devices are identical one could be materially random and the other not. The causal anti-realist can deny that there are any *essentially* random devices. He can argue that whether the device is random depends not on the nature of the device but on various events and regularities. He can argue that his forwards-causal description of the situation shows that the device is not random. *In the third case*, where the device is statistically random, assume that there are a thousand patterns from which X is selected. Leaving aside, for the moment, the question of whether the device is random, assume that the precogniser affects the device in such a way that whichever of the thousand patterns he draws that pattern is subsequently displayed. It follows that the patterns displayed vary over the thousand in precisely the same manner as the drawings. Therefore if the precogniser draws at random one of the thousand patterns, it is *necessarily* the case that the patterns are subsequently selected in a statistically random fashion. Consequently the statistical randomness of the device for selecting the pattern is consistent with the precogniser affecting the random device.

Objection Two: The description of the situation in forwards-causal terms involves an excessively complicated hypothesis, because:

(1) a device which is usually *uncaused* is on this

occasion *caused*;

(2) the precognition affects the "random" numbers generator, but the precogniser knows neither what the numbers are nor the procedure for converting random numbers into patterns.

To this objection, I reply that the complexity of the hypothesis is entirely due to the insistence that it be stated in *causal terms* - I do not deny that precognition has occurred, and I do not deny that the content of the precogniser's knowledge has nothing to do with random numbers. The situation can be simply described as one in which it is nomically necessary and sufficient that the precogniser draw the pattern which is displayed the next day. No causal description is required. But the causal anti-realist can, if he wishes, interpret the situation as one which involves forwards causation (See Section 3.1).

3.4 On Clairvoyance and Precognition

Consider a precognition experiment like that described in Section 3.3, but in which before the precognition the random numbers are generated which are used to select a pattern after the precognition. In this case, it is already determined what pattern will be displayed. The description of the situation in terms of the laws of Nature has not altered; it is *nomically necessary* (given the state of the precogniser's mind) that the precogniser draws a certain pattern if and only if that pattern is displayed a day later. But the *causal description* designed to avoid backwards causation is different. One *now* says that the precognition and the display of the pattern are both caused by the random selection of numbers. It is important not to confuse such a case of *precognition* with *clairvoyance*. If a clairvoyant knew the random numbers, and if he also knew the procedure for converting the

numbers into patterns he could then *infer* that a certain pattern would be displayed. But the precogniser need not know anything about either random numbers or selection procedures; the content of his knowledge is the display of the pattern a day later. Nor am I suggesting that he is an *unconscious clairvoyant*; I am simply describing, in a complicated causal fashion, what can be more simply described in terms of nomic necessity.

Finally, suppose one sometimes selects the random numbers before, and sometimes after, precognition. In that case the situation involving the same laws of Nature has to be variously interpreted as one of the precognition affecting the random device and as a case of collateral effects. Such complexity, while it might embarrass the causal realist, does not embarrass the anti-realist who has, in all cases, a simple description in terms of laws of Nature.

The discussion in Sections 3.2, 3.3 and 3.4 shows, I claim, that the causal anti-realist need not treat hypothetical examples of precognition as cases of backwards causation. Furthermore, the distinction between precognition and clairvoyance is one of the *content* of the knowledge (conscious or unconscious) rather than one of the kind of causal relations involved.

3.5 Why do Causes Explain their Effects, but not Vice Versa?

An objection might be raised to Causal Anti-realism on the grounds that without some irreducible element of causal priority "in the objects", it is totally mysterious why effects do not explain causes. I attempt to answer this objection without discussing the various theories of explanation - such as the Covering Law Theory - which would be beyond the scope of this thesis.

I begin by distinguishing between absolute and relative explanations.

An *absolute* (but perhaps *partial*) explanation of some state-of-affairs is an account which makes the state-of-affairs either simpler or less mysterious. A *relative* explanation of one state-of-affairs in terms of another is some account of why, given the explanans the explanandum is simpler or not so mysterious as it would otherwise be. In other words, in an absolute explanation the explanandum together with the explanans is simpler or less mysterious than the explandum by itself. Whereas in a relative explanation the following counterfactual is correct:

> If the explanans were to be treated as simple or as not mysterious then pointing to the explanans renders the explandum simpler or less mysterious than it would otherwise seem to be.

For example, suppose that one held a "Big Bang" theory in which the initial state of the universe is very simple, and suppose one could show that given this initial state and the laws of Nature, there would *have to* evolve the complex arrangement of galaxies in clusters which astronomers observe. In that case, one would have provided a partial, absolute, explanation of the arrangement of galaxies in clusters. However, suppose one explained the occurrence of ancient paintings on an uninhabitable island near Antarctica in terms of this island having been inhabited 2,000 years ago. Unless one could explain why this island was then inhabited, this explanation would not be absolute: For, in this case the conjunction of the explanandum and the explanans is quite as mysterious as the explanandum by itself.

Note: A *mark* of the difference between absolute and relative explanations is that if p would, if it were true, be an *absolute* (but perhaps partial) explanation of a known truth q then one has grounds for inferring by *abduction* that p is true, whereas one should not use

abduction to argue to a purely relative explanans. For instance, if one argued rationally to the past inhabitation of the island from the presence of paintings on the island, the argument would have to be by ordinary *induction* not by *abduction*.

In the case of causation, the occurrence of the cause and of the effect can be, I suggest, partial or total *relative explanations of each other*. The house burning is explained in terms of the fault in the wiring, but one can also provide what von Wright calls *quasi-teleological explanations*:

> We observe that the state following after the explanandum c is the state d; c we think, is a necessary condition of this state. The state d materialised - but had it not been for c, d would not have come about; c was needed to make it possible, one could say. We are are here interested in explaining d. We take its occurrence for granted. [24]

For example, I observe the behaviour of a mechanical tortoise, and I am puzzled until I realise that its complicated behaviour is necessary in the circumstances for it to recharge itself from an electric power-point. I have then provided a relative explanation of the behaviour of the tortoise in terms of its recharging itself; that is, I have explained the cause in terms of the effect. One is tempted to say that the explanation is in terms of either the tortoise being designed to recharge itself or having some goal-seeking mechanism. But I suggest that the *relative* explanation would still hold if I did not know whether the tortoise was designed or not, or *even if* I discovered that the tortoise only contained a battery and motor and no complicated electronics, provided the tortoise could not recharge itself without moving roughly as it did. If the tortoise only contained a battery and motor, this quasi-teleological explanation would

---

24. von Wright, G.H., *Explanation and Understanding*, pp. 57,58.

not be an absolute explanation.

The regularity-theorist can give an account of relative causal explanations; instances of a regularity seem less mysterious than pairs of events that cannot be subsumed under a regularity. Notice that if one is considering relative explanations, one can also explain one collateral effect in terms of another provided the two are constantly conjoined. For example, one could explain the alarm (which is set for 7 o'clock) going off at 5 o'clock in terms of the clock hands being at the 7 o'clock position, even though one might be totally ignorant of whether the position of the clock hands and the alarm going off are cause and effect or collateral effects.

Now consider *absolute explanation*. I suggest that there are two reasons why causes seem to explain effects but not vice versa.
(1) If an agent acts (event X) in order to produce Y, the explanation of the agent's action X is his intention to produce Y. So there is a relative explanation of Y in terms of X which is itself explained in terms of the intention to produce Y. That an agent who has what are generally considered to be good reasons to act as well as the power to act, does act, might be considered a state of affairs *lacking in mystery*. So one obtains a partial but an *absolute* explanation of the occurrence of Y, and this explanation involves a *chain of relative explanations* from the agent's intention to X to Y. I suggest that one of the reasons for treating causes as partial absolute explanations of their effects is that often the cause itself can be explained in terms of an agent's actions.

It is interesting to notice that in the case where the agent *seems* to produce the effect in order to produce the cause (for instance he raises his arm in order that certain neural events occur earlier)

one *seems* to provide a partial absolute explanation of the mysterious cause (the neural event) in terms of the less mysterious effect (the raising of the arm). Here the quasi-teleological explanation is apparently *absolute*, not relative.

(2) One can consider causal chains in which C causing E forms one link. It is logically necessary that, on the whole, the chains stretch from past to future because otherwise there would, I suggest, be no coherent temporal ordering.[25] Irrationally perhaps, it is assumed that what occurs in the distant past (or at the initial moment) is peculiarly lacking in mystery -- it is to be taken for granted. By tracing a causal chain backwards from effect to cause one approaches the distant past, but if one traced the chain from cause to effect one would approach the distant future which is thought of as mysterious. Hence explanations of effect in terms of causes (even if the causes are later than the effects) would be relative, not absolute, for they are leading one in the *wrong direction along the causal chain.*

The justification, if any, for this bias in favour of the past over the future is not any supposed simplicity of the initial state of the universe, for the initial state could be complicated and the final state very simple. I suggest - speculatively and tentatively - that this bias has two sources. The first source might be the influence of crude ideas about creation: God acts like a human being and so produces *the cause* (the initial state of the universe) *in order to produce the effect.* The second source might be an obscure metaphysical theory that in some sense the past, being fixed, has *ontological priority* over the future and so is less mysterious than the future. It follows that the causal anti-realist might be committed to denying

---

25. See my account of time (Section 1.5), pp. 191-194.

widely held views that the past explains the future, and consequently denying that causes *absolutely* explain effects. Nonetheless, he can sketch an account of why causes *seem* to explain effects, but effects do not seem to explain causes.

3.6 The Interpretation of Situations in which, it is Claimed, Several Earlier Events are the Collateral Effects of a Later Event.

Suppose quasi-events $E_1 \ldots E_n$ occur and later a quasi-event L occurs, and the co-occurrence of $E_1 \ldots E_n$ can be explained if one accepts that L causes $E_1 \ldots E_n$. The quasi-events $E_1 \ldots E_n$ might be the various drawings by different precognisers. L might be the display of a randomly selected pattern. The challenge to produce an *alternative simple explanation* in terms of *forwards causation* is *misguided*. For, as I argue, there is a satisfactory explanation either in terms of laws of Nature or in terms of regularities. If one *demands* a forwards causal description, one can be provided but it is very complicated.

Given the precognisers' state of mind, if it is nomically necessary that the pattern which they draw is later displayed, then it is also nomically necessary that they agree with one another. Hence, assuming that only one pattern is to be displayed, the co-occurrence of the events does not require a causal explanation at all; it only requires there to be laws which ensure that each drawing is sufficient for the display of the patterns. That the precognisers agree is a consequence of their ability to precognise; it does not require any explanation in causal terms. A roughly analogous situation involving overdetermination is the case in which several mechanical tortoises

all find the one power-point in the room and the first to get there switches on that power-point. The behaviour of any of the tortoises is sufficient in the circumstances for the power-point to be switched on, and because there is only one point, it is not a coincidence that all tortoises go to the same place. Furthermore, suppose there is no mechanism of any kind inside the tortoises except a battery and a motor yet it is a law of Nature that mechanical tortoises move towards electric power-points when their batteries have run down. In that case although this law explains the tortoises all going to the same point, the situation might nonetheless be interpreted as one in which a *quasi-teleological* explanation rather than a backwards causal explanation is appropriate. The case of the precognisers' drawing the correct pattern has a somewhat more complicated forwards causal description. For, if before the precognition, one had taken steps to *prevent* the pattern being displayed, then the events $E_1 \ldots E_n$ would *not* have occurred, but if one had fixed the switch in the off position, the tortoises *would* nonetheless have moved towards the power-point. Conversely, if one were to prevent all the precognisers' drawing, one would *not* prevent *some* pattern being displayed. But in the circumstances, if one were to prevent all the mechanical tortoises from moving one would prevent the power-point being switched on. Also, in the precognition example, the pattern which is displayed is selected on a statistically random basis.

In the example of the several precognisers there are interlocking (forwards) causal and logical relations.

(1) Some quasi-event D_1 prevents interference by any agent.[26]

(2) Some quasi-event D_2, perhaps the same as D_1, causes there to

---

26. See p. 181, where I discuss Black's argument.

be a limited range of possibilities in the display of the pattern. The effects of D_2 are:

(i) that one and only one pattern is displayed.

(ii) that this pattern is chosen from a given range of patterns.

Let R be the conjunction of these two effects.

(3) D_1, D_2 or some other quasi-event is necessary but not sufficient in the circumstances for E_1, ..., E_n (the quasi-events that the precognisers make their drawings).

(4) The ith precogniser's drawing the pattern X_i causes the display of the pattern X_i. (Drawing X_i is sufficient in the circumstances for the display of X_i.)

(5) The occurrence of R *logically entails* that all the patterns X_i which are displayed are the same.

Note One: R is not causally effective since R *entails* that all the X_i to be displayed are the same. The *cause* of the patterns all being the same is the quasi-event D_2. Hence D_2 *forces* the precognisers to agree, just as there being only one power-point *forces* the mechanical tortoises to move to the same point.

Note Two: Given the circumstances before the precognition, D_2 but not E_i is necessary for the selection of the pattern drawn by the ith precogniser. (E_i is not necessary since even if E_i had not occurred R would ensure that *some* pattern is selected so it might have been the pattern which, in fact, the ith precogniser drew.)

As I warned the description is of some complexity, but this complexity, I claim, does not embarrass the causal anti-realist. There are, however, four other objections to this account to which I now reply.

Objection One: The above account is only a *relative* explanation of the agreement of the precognisers, but for an absolute explanation one requires backwards causation.

Reply: The agreement of the precognisers is explained in terms of the laws of Nature and the quasi-event D_2, which causes only one pattern to be displayed. A backwards causal explanation would also be in terms of one and only one pattern being displayed. The success of an *absolute* (but perhaps partial) explanation depends in both cases on how one explains the occurrence of D. Notice that if *several* patterns were *displayed*, but the precogniser *drew only one pattern*, even if one were using the notion of backwards causation one would be driven to the conclusion that there was some interaction between the precognisers; there being only one pattern is an essential feature of the backwards causal explanation.

Objection Two: It is inconceivable that the disjunctive state-of-affairs R is fixed at a time (the occurrence of D_2) when none of the disjuncts is fixed.

Reply: Consider the following hypothetical example. It is discovered that if ordinary lead is exposed to a beam of α-rays, the lead, without any other apparent change, becomes radio-active and inevitably decays within two months. There are two possible interpretations. The first is to insist that some undetectable change has occured in the lead. The second is to say that it is nomically necessary that when a lead atom has been bombarded by an α-ray it will decay within two months. Now one might insist on the former interpretation, but the latter is, surely, a *coherent* interpretation. If one wished one could then say that bombarding the lead by an α-ray (mnemically) causes the indeterminate state of affairs that the lead decays at some time in the two months.

If one rejects any causal description in this case - as one might - one should not, I suggest, demand a causal description of the fixing of R by D_2. One should simply say that it is *nomically necessary* that if D_2 occurs then some pattern appears.

Objection Three: If the precognisers affect the selection of the patterns then several precognisers have the power to freely decide whether to bring about the *quantitatively* identical event, which is absurd.

Reply: The precognition would be the *producing cause* of the selection of the pattern. It does not follow that the precogniser has *power* over the pattern. The precogniser is not a fraud with extra-sensory powers other than precogntion; he is a genuine precogniser. As such he has no *power* over what he precognises, even though his precognition *causes* the pattern to be selected.

Objection Four: Suppose a random device is first used to select the number of patterns to be displayed and each pattern is subsequently selected at random. If it then happens that only one pattern is selected, one can no longer explain the agreement of the precognisers' drawing in terms of forward causation.

Reply: *On the one hand* there is still a forwards causal *relative* explanation; if only one drawing is selected, the precognisers' drawings agree (for unless the precognisers' drawings agree there would be more than one drawing). But, *on the other hand*, the backwards causal explanation is also *no longer absolute* for one cannot explain why only one pattern was displayed.

No suppose that it is rather unlikely that more than one pattern is displayed. In that case, Black's argument[27] shows that there is

27. See p.181.

some quasi-event D_1 no later than the precognition such that D_1 prevents interference with the number of patterns shown. So, although D_1 is *not sufficient* in the circumstances for R, D_1 *renders R likely;* consequently the forwards causal explanation is as successful as the backwards causal explanation is in explaining why the precognisers agree.

Section Four: A Discussion of Some Proposed Hypothetical Examples of Backwards Causation

In this section I discuss five hypothetical examples which one is tempted to treat as cases of backwards causation. In all cases I argue that a *causal anti-realist* need not treat the situation as one which involves backwards causation.

4.1 Three Preliminary Points of Clarification

(1) I am not claiming that a causal anti-realist is committed to denying that backwards causation occurs. I am claiming that no hypothetical example could force him to accept backwards causation on pain of irrationality.

(2) If some of the situations described actually occurred, people might begin to call them cases of backwards causation. But, whether they would or not is of no consequence to the causal anti-realist. The situations are describable without using the notion of a producing cause, nor would there be any good reason for closing the open texture of the notion of cause in this respect.

(3) I assume throughout this section that peoples' intuitions are divided on the issue of backwards causation, and so one cannot simply point to an example as an accepted case of backwards causation. On

the one hand, one intuitively tends to deny that a cause could be later than its effect. On the other hand, one intuitively tends to accept Causal Realism and the causal realist has, I shall argue, grounds for describing some hypothetical examples as examples of backwards causation.

4.2 Example One - Precognition

In Section 3.2 I argue that in some cases of precognition, the knowledge could cause the known event. In section 3.4 I argue that in other cases of precognition the knowledge and the known event could be treated as collateral effects without treating the precogniser as a very clever clairvoyant who merely pretends to be a precogniser. The apparent difficulty that the precognised event might be randomly selected is answered by distinguishing between a kind of event whose occurrence is statistically random, and an event without a cause. Except by begging the question, one cannot assume that the device is random in the sense of having no cause. But if the precognisers' drawing is itself statistically random, and it affects the device so that there is a one to one correspondence between the pattern drawn and the pattern selected, then the device not only can be but also must be itself statistically random. In order to avoid a backwards causal description of precognition, an extremely complicated description is required. Depending on when the pattern was selected, the situation is sometimes described as one of *collateral effects*, and sometimes as one of *cause and effect*. Sometimes the device is random only in the statistical sense; sometimes it is random also in the sense of having no cause. But as I argued in Section 3.1, the complexity of the description need not embarrass the causal anti-realist. However, the causal realist might interpret this complexity as grounds for believing in the possibility of backwards causation.

4.3 Example Two: An Agent Affecting the Past

If an agent intentionally performs action A in order to bring about an earlier event B, and whenever he so acts the earlier event has occurred, one might be tempted to describe the situation as one in which A causes B. Consider again Dummett's example mentioned in Chapter One:

> Someone who believes in magic ... has ..., among his spells a formula for producing good weather in a particular place on a particular day ... An occasion arises when he has reason for wanting the weather at, say, Liverpool, to have been good on the previous day, but he does not know whether there was [good weather] or not; he therefore recites his spell, putting in yesterday's date. Subsequently he finds out that there is fine weather at Liverpool on that day; ... whenever he recites the formula with a past date, not knowing what the weather was like ..., later investigation proves the weather to have been fine then. [28]

In this case Gale's argument[29] shows that it was fixed, at the time the good weather occurred in Liverpool, that the magician would not know that good weather had occurred when he recited the spell. Also, as Black's (and perhaps Swinburne's) argument shows, although the reciting of the spell was intentional it would be free only in a *compatabilist* sense, and since the action is intentional it is not random: hence it is determined by some previous event D. With these provisos, the situation seems coherent. The argument that it should be considered a case of backwards causation is based on its similarity to the paradigms in which an agent brings A about in order to bring B about.

But now consider again Chisholm's example of the man raising

---

28. A.E. Dummett, "Can an Effect Precede its Cause?' *Aristotelian Society*, supp.vol.28. pp. 35,36.

29. See p. 187.

his arm in order to bring about some earlier neural event. Unless one accepts *that* example as a case of *backwards* causation it shows that the intended *means* to an end can be the *effect* of the intended *end*. Likewise, in the example of the magician one might argue that the intended means is not the cause of the intended end. In this case the intended end and the intended means to that end are collateral effects. For some quasi-event D, say the state of mind of the magician, determines that the magician recites his spell and so, even on the backwards causal hypothesis, D would be an (indirect) cause of the fine weather. The causal anti-realist treats the situation as a case of collateral effects rather than as a causal chain.

Note: If one does not accept the corollary to Black's argument[30] (that is, no agent *freely* affects the past, in a libertarian sense of 'freely'), one might interpret the situation as one of backwards *immanent* causation. The magician *directly* affects the weather, and the fine weather causes the magician to recite his spell. Here the intended *means* is the *effect* of the intended *end*. Alternatively, someone might argue that some event in the magician's mind - call it a volition - is itself uncaused and causes the fine weather. In that case the interpretation would indeed involve backwards *transeunt* causation. But if one analyses a free action as an action which has as its cause a volition which itself has no cause, then one has given a description of the proposed example in transeunt causal terms; by saying that the action is free one has already stipulated that the volition is not to be treated as an effect. Hence one has not provided an example described in *non-causal* (that is, not transeunt causal) terms which is to be *interpreted* as a case of backwards causation.

---

30. See p. 184.

4.4 Example Three - "Time-reversal" Examples

Suppose that in some distant part of the universe there is a planet on which, it seems, the temporal ordering is reversed. In other words, the processes occurring on that planet seem quite unlike those occurring on Earth, until it is noticed that they are the temporal inverses of processes very like those occurring on Earth. Thus one is to suppose that on that planet entropy is apparently decreasing.[31] Someone might argue that both on Earth and on that planet lightning strikes cause bush fires, not vice versa. Hence, *either* on Earth *or* on that planet, there is a case of *backwards causation*.

I call these cases "time-reversal" examples because one is tempted to say that time is "flowing in different directions" in different parts of the universe, so there is no backwards causation. However, one cannot coherently say that time flows in different directions in different parts of the universe. For either the two parts are spatially disconnected, in which case one cannot even *compare* the direction of their temporal orderings, or the two parts are spatially connected so there is a boundary between the two parts consisting of at least one point. In order that the temporal orderings in the two parts of the universe be compared, there must be some correlation of events at the boundary. So one may assume that A_1 and A_2 are simultaneous; A_1 is on one side and A_2 on the other side of the boundary. Likewise B_1 and B_2 are simultaneous; B_1 is on the same side as A_1, B_2 is on the same side as A_2. But if A_1 is earlier than B_1 but A_2 is later than B_2, the temporal ordering cannot be transitive for A_2 is simultaneous with A_1 which is earlier than B_1 which is simultaneous with B_2, so by transitivity A_2 is earlier than B_2. But A_2 is later than B_2.

Because one cannot describe the world as having two different

---

31. This kind of example is due to Michael Tooley.

temporal directions (except by sacrificing transitivity) one might be tempted to interpret the situation as involving backwards causation in one part or the other. However, I suggest that the ordinary temporal ordering, which is necessarily transitive, presupposes the substantial agreement of the proto-temporal orderings of phenomenological time, physical time, and causal priority. So in such a case *there is no coherent temporal ordering* and hence no backwards causation. There would however still be *local* temporal ordering which would apply only to the different parts of the universe.

4.5 Example Four

Suppose that on the vast majority of occasions in which one heats an iron bar to 1000° the bar glows. Now suppose that, very rarely, the iron begins to glow after the bar is heated, and, even more rarely, the iron begins to glow *before* the bar is heated.

In that case the simplest causal hypothesis is that heating the iron bar *always* causes it to glow. Hence, if one requires a simple causal hypothesis, one is committed to there being some cases of backwards causation. This kind of example can be elaborated. Suppose that *usually* events of kind C are followed by events of kinds E_1, E_2, ..., E_n, but that very rarely one or more of the E_1, ..., E_n precede the C. The forwards causal hypothesis is a very complicated disjunction of the cases:

The C causes the E_1 ... E_n if the C is the earliest event;

The E_1 causes the C, E_2 ... E_n if the E is the earliest event;

Etc.

The simplest causal hypothesis is that in all cases the C causes the E_1, E_2, ..., E_n. However, in such cases there is a simple description

in terms of laws of Nature: 'If a *C* occurs, then shortly before or shortly after some E_i occurs' for i=1 ... n. As I have said in Section 3.1, the complexity of an unnecessary forwards causal explanatory hypothesis does not embarrass the causal anti-realist. However, the realist about causal priority would have grounds for interpreting this situation as one which involves backward causation.

4.6 Example Five - Collateral Effect Examples

In some situations one might be able to explain the co-occurrence of quasi-events E_1, ..., E_n on the hypothesis that they are all collateral effects of some later cause C (of kind *C*). I have already discussed this in Section 3.6 using the example of the several precognisers who all draw the same pattern. There is no need whatever for a causal explanation; the co-occurrence of E_1, ..., E_n is explained just as readily by assuming that each of the E_i is *nomically sufficient* for some event of kind *C*, and by assuming that it is already fixed that only one event of kind *C* occurs.

In this case one might say that C is the explanation of E_1, ..., E_n, but C need not be interpreted as the *producing* cause of E_1, ..., E_n. I suggest that a tendency to assimilate explanations to producing causes lies behind the urge to treat this kind of example as one of backwards causation. But, as in the example of the mechanical tortoise, there seem to be cases of *quasi-teleological* explanations where the effect explains the coincidence of several overdetermining factors. As in the previous examples, the complexity of the forwards causal description need not embarrass the causal anti-realist, but it provides some grounds for the causal realist to accept the possibility of backwards causation.

4.7

The discussion of these five examples of proposed cases of backwards causation illustrates the general strategy that a causal anti-realist could use, if he proposed an analysis of causation incompatible with the occurrence of backwards causation. He provides suitably simple descriptions of the situation in terms of laws of Nature and, if one is demanded, he provides a complicated description in forwards-causal terms.

BIBLIOGRAPHY

Anderson, J., "The Problem of Causality", *Australasian Journal of Psychology and Philosophy*, vol.16 (August 1938), pp. 127-142.

Anscombe, G.E.M., "Causality and Determination" in *Causation and Conditionals*, ed. Sosa, E., London, New York, Oxford University Press, 1975, pp. 63-81.

Beauchamp, T.L. (ed.), *Philosophical Problems of Causation*, Encino, Calif., Dickenson Pub. Co., 1974.

Berkeley, G.A. *A Treatise Concerning the Principles of Human Knowledge*, Reprinted ed. Chicago, Open Court, 1901.

Black, M., "Why Cannot an Effect Precede its Cause?' *Analysis* vol.16 (Jan. 1956), pp. 49-58.

Brier, B., *Precognition and the Philosophy of Science; an Essay on Backward Causation*, New York, Humanities Press, 1974.

Chisholm, R.M., "Freedom and Action" in *Freedom and Determinism*, ed. Lehrer, K., New York, Random House, 1966, pp. 11-44.

Davidson, D., "Causal Relations", *The Journal of Philosophy*, vol 64 (Nov. 1967), pp. 691-703.

Descartes, R., *Philosophical Writings: A Selection*, trans. and ed. Anscombe, G.E.M. and Geach, P.T., Melbourne, Nelson, 1966.

Ducasse, C.J., *Truth, Knowledge and Causation*, London, Routledge and Kegan Paul, New York, Humanities Press, 1968.

Dummett, A.E., "Can an Effect Precede its Cause?", *Aristotelian Society*, Supplementary Volume 28 (July 1954), pp. 27-44.

Flew, A., "Can an Effect Precede its Cause?" *Aristotelian Society*, Supplementary Volume 28 (July 1954), pp. 45-62.

_____, "Causal Disorder Again", *Analysis*, vol.17 (March 1957), pp. 81-86.

Gale, R.M. "Why a Cause Cannot be Later than its Effect", *The Review of Metaphysics*, vol.19 (Dec.1965), pp. 209-234.

Gasking, D., "Causation and Recipes", *Mind*, vol.64 (October 1955), pp. 479-487.

Goodman, N., *Fact, Fiction and Forecast*, 3rd ed., Indianapolis, Bobbs-Merrill, 1973.

Hughes, G.E. and Cresswell, M.J., *An Introduction to Modal Logic*, London, Methuen, 1968.

Hume, D., *A Treatise of Human Nature*, ed. Selby-Bigge, L.A., Oxford, The Clarendon Press, 1896.

_____, *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals*, ed. Selby-Bigge, L.A., 2nd ed. Oxford, The Clarendon Press, 1902,

Kant, I., *Critique of Pure Reason*, trans. by Kemp Smith, N. London, Macmillan, 1929.

Kim, J., "Causes and Events: Mackie on Causation", *Journal of Philosophy*, vol.68 (July 1971), pp. 426-441.

_____, "Causes and Counterfactuals", *Journal of Philosophy*, vol. 70 (October 1973), pp. 570-572.

Lehrer, K. (ed.), *Freedom and Determinism*, New York, Random House, 1966.

Lewis, D.K., *Counterfactuals*, Oxford, Blackwell, 1973.

_____ , "Causation", *Journal of Philosophy*, vol.70 (October 1973), pp. 556-567.

Lyon, A., "Causality", *British Journal of the Philosophy of Science*, vol.18 (1967), pp.1-20.

Mackie, J.L., "Counterfactuals and Causal Laws", in *Analytical Philosophy*, ed. Butler, R.J., Oxford, Blackwell, 1962, pp. 66-80.

_____, *Truth, Probability and Paradox: Studies in Philosophical Logic*, Oxford, Clarendon Press, 1973.

_____, *The Cement of the Universe: A Study in Causation*, Oxford, Clarendon Press, 1974.

Martin, R., "The Sufficiency Thesis", *Philosophical Studies*, vol. 23 (1972), pp. 205-211.

Mill, J.S. *A System of Logic; Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation*, 6th ed., London, Longmans, 1865.

Montgomery, H.A. and Routley, F.R., "Contingency and Non-contingency Bases for Normal Modal Logics", *Logique et Analyse*, vol.9 (1966), pp. 318-328.

Nagel, E., *The Structure of Science: Problems in the Logic of Scientific Explanation*, London, Routledge, 1961.

Pears, D.F., "The Priority of Causes", *Analysis*, vol. 17 (Jan.1957), pp. 54-63.

Quine, W.V.O., "On What There Is" in *From a Logical Point of View; Nine Logico-philosophical Essays*, 2nd ed. rev., Cambridge, Harvard University, 1953, pp. 1-19.

_____, "Two Dogmas of Empiricism" in *From a Logical Point of View*, pp. 20-46.

_____, *The Roots of Reference*, La Salle, Ill., Open Court, 1973.

Reichenbach, H., *Nomological Statements and Admissible Operations*, Amsterdam, North-Holland Pub. Co., 1954.

Scriven, M., "Randomness and the Causal Order", *Analysis*, vol.17 (Oct. 1956), pp. 5-9.

Smart, J.J.C., *Philosophy and Scientific Realism*, London, Routledge and Kegan Paul, 1963.

Sober, E., *Simplicity*, Oxford, Clarendon Press, 1975.

Sosa, E (ed.), *Causation and Conditionals*, London, New York, Oxford University Press, 1975.

Swinburne, R., *Space and Time*, London, Melbourne, etc., Macmillan; New York, St. Martin's Press, 1968.

Taylor, R., *Action and Purpose*, Englewood Cliffs, N.J., Prentice Hall, 1966.

von Wright, G.H., *Explanation and Understanding*, London, Routledge and Kegan Paul, 1971.

_____, *Causality and Determinism*, New York, Columbia University Press, 1974.