



THE UNIVERSITY OF ARIZONA

University Libraries

UA CAMPUS
REPOSITORY

PROBABILITY AND CAUSALITY.

Item Type	text; Dissertation-Reproduction (electronic)
Authors	OTTE, RICHARD EDWARD.
Publisher	The University of Arizona.
Rights	Copyright © is held by the author. Digital access to this material is made possible by the University Libraries, University of Arizona. Further transmission, reproduction or presentation (such as public display or performance) of protected items is prohibited except with permission of the author.
Download date	21/10/2021 15:52:46
Link to Item	http://hdl.handle.net/10150/187748

INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.
2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again--beginning below the first row and continuing on until complete.
4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB RD., ANN ARBOR, MI 48106

8217449

Otte, Richard Edward

PROBABILITY AND CAUSALITY

The University of Arizona

PH.D. 1982

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

Copyright 1982

by

Otte, Richard Edward

All Rights Reserved

PROBABILITY AND CAUSALITY

by

Richard Edward Otte

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF PHILOSOPHY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 8 2

© Copyright 1982 Richard Edward Otte

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read
the dissertation prepared by Richard Edward Otte
entitled Probability and Causality

and recommend that it be accepted as fulfilling the dissertation requirement
for the Degree of Doctor of Philosophy.

Wesley C. Salmon 4/22/82
Date
Henry Byerly 4/22/82
Date
John D. Clark 4/22/82
Date

Date

Final approval and acceptance of this dissertation is contingent upon the
candidate's submission of the final copy of the dissertation to the Graduate
College.

I hereby certify that I have read this dissertation prepared under my
direction and recommend that it be accepted as fulfilling the dissertation
requirement.

Wesley C. Salmon 4/22/82
Dissertation Director Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation form or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

Signed Richard Edward Otto

PREFACE

I would like to thank the D. Reidel Publishing Company for allowing me to include a previously published paper (Otte 1981) in this dissertation. All of the material in chapter 4 and the beginning of chapter 6 falls into this category. Their cooperation on this matter is appreciated.

Although I would like to express my gratitude to everyone who over the years encouraged me in my academic endeavors, the list would be far too long. I would first like to thank my parents for instilling within me a love of knowledge and a desire to learn; their encouragement played a major role in my pursuing an advanced degree. Numerous other friends provided encouragement when I felt the discouragement typical to most graduate students; to them I extend my gratitude. I would also like to thank the various faculty members who have helped shape me as a philosopher. Their patience in training me to be a philosopher is deeply appreciated.

Finally, I wish to express my gratitude to Professor Wesley C. Salmon. One could not have hoped for a better dissertation director. I was truly fortunate to have been able to work with him, and my indebtedness to him will be apparent as one reads this dissertation.

TABLE OF CONTENTS

1. PROBABILISTIC CAUSALITY	1
1.1. The Problem of Identifying Deterministic Causes	1
1.1.1. Hume's Need For Probabilistic Causes	2
1.1.2. Mackie's Account of Causation	3
1.2. The Problem of Ordinary Causal Concepts	4
1.2.1. Suppes' Examples	4
1.2.2. A Response to Suppes	4
1.2.3. Suppes and Epistemic Causal Concepts	5
1.3. Reichenbach and Indeterminism	5
1.4. Salmon and Probabilistic Causality	6
1.4.1. Salmon's Argument	6
1.4.2. A Reply to Salmon	6
1.4.3. Salmon's Rejoinder	7
1.5. Modern Physics and Indeterminism	7
1.5.1. Causation and Indeterminism	7
1.5.2. Examples of Indeterministic Causation	8
1.6. Positive Relevance	9
2. HISTORICAL ROOTS OF PROBABILISTIC CAUSALITY	11
2.1. Hume's Discussion of Causality	11
2.1.1. Hume's Theory of Deterministic Causality	11
2.1.2. Hume's Theory of Probability	15
2.1.3. Hume's Discussion of Probabilistic Causation	18
2.1.4. Critical Discussion of Hume's Theory of Probabilistic Causation	21
2.1.5. Concluding Remarks on Hume	28
2.2. Mill's Theory of Causality	29
2.2.1. Mill's Defense of Deterministic Causality	30
2.2.2. Mill's Methods and Deterministic Causality	31
2.2.3. Modification of Mill's Methods	32
2.3. Gasking's Treatment of Causality	33
2.3.1. Brief Statement of Gasking's Theory	33
2.3.2. Modification of Gasking's Theory	34
3. REICHENBACH'S TREATMENT OF CAUSALITY	35
3.1. The Relation of Causal Betweenness	35
3.1.1. Definition of Causal Betweenness	35
3.1.2. Explication of Definition 1	36
3.1.2.1. Positive Relevance.	36
3.1.2.2. Screening Off.	37
3.1.2.3. Causal Betweenness and the Causal Net.	37
3.1.3. Critical Discussion of Causal Betweenness	38

v	
3.2. Causal Forks	39
3.2.1. Conjunctive Forks	39
3.2.2. Conjunctive Forks and Common Causes	40
3.2.3. Interactive Forks	41
3.3. Causal Relevance	43
3.3.1. Causal Relevance and Mark Transmission	43
3.3.2. Causal Relevance and Probability Relations	43
3.3.2.1. Assumption Alpha.	43
3.3.2.2. Assumption Beta.	44
3.3.2.3. Assumption Gamma.	44
3.3.2.4. Assumption Delta.	45
3.3.2.5. Probabilistic Definitions of Causal Relevance.	45
3.3.3. Causal Betweenness, Conjunctive Forks, and Causal Relevance	45
3.3.4. The Equivalence of Definitions 2 and 4	46
3.3.5. Modifications of Assumption Gamma	54
3.3.6. Problems With Definition 3	56
3.3.6.1. The Occurrence of The Cause and Effect.	56
3.3.6.2. Suppes' Criticism of Definition 4.	59
4. SUPPES' PROBABILISTIC THEORY OF CAUSALITY	61
4.1. Exposition of Suppes' Theory	62
4.1.1. Prima Facie Causes	62
4.1.2. Spurious Causes	62
4.1.3. Direct and Indirect Causes	64
4.1.4. Supplementary Causes	65
4.1.5. Sufficient Causes	66
4.1.6. Necessary Causes	66
4.2. Analysis of Spurious ₁ and Direct ₁ Causes	67
4.2.1. Example 1	67
4.2.2. Example 2	69
4.2.3. Example 3	71
4.3. Analysis of Spurious ₂ and Direct ₂ Causes	72
4.3.1. Examples 2 and 3	72
4.3.2. Example 1	73
4.3.3. Example 4	73
4.3.4. Suppes' Defense of the Requirements of Spurious ₂	74
4.4. Spurious ₃ and Direct ₃ Causes	76
4.4.1. The Definition of a Spurious ₃ Cause	76
4.4.2. The Definition of a Direct ₃ Cause	77
4.4.3. Relations Among the Definitions of Spurious Causes	77
4.4.4. Spurious ₃ Causes and Examples 2, 3, and 4	79
4.4.5. Spurious ₃ Causes and Example 1	79
4.5. Spurious Causes and Interactive Forks	81
4.6. Final Remarks on Suppes	81

5. GOOD'S QUANTITATIVE CAUSAL CALCULUS	83
5.1. Two Different Causal Concepts	83
5.1.1. The meaning of $\text{Chi}(E:F)$	84
5.1.2. The Meaning of $Q(E:F)$	85
5.1.3. $\text{Chi}(E:F)$ and $Q(E:F)$	85
5.2. Good and Spurious Correlations	86
5.2.1. Salmon's Objection	86
5.2.2. Good's reply to Salmon	86
5.2.3. Discussion of Good's reply	87
5.2.4. Partially Spurious Correlations	89
5.3. The Strength of Causal Chains	91
5.3.1. Salmon's Counterexample	91
5.3.2. Good's reply	94
5.4. Quantitative Theories of Causality	95
6. THE QUESTION OF POSITIVE RELEVANCE	96
6.1. Prima Facie Causes and Negative Relevance	96
6.1.1. Hesslow's Argument	96
6.1.2. Probabilistic Analogue of a Causal Field	97
6.2. Rosen's Defense of Positive Relevance	98
6.2.1. Rosen and Adequate Conceptual Frameworks	98
6.2.2. Rosen and Improbable Consequences	102
6.2.3. Rosen's Original Example	103
6.2.4. Objectively Homogeneous Reference Classes and the Precise Specification of Events	104
6.3. Indeterminism and the Improbable	106
6.4. I. J. Good and Positive Relevance	107
6.4.1. Good's Method of Interpolated Causal Links	107
6.4.2. Salmon's Counterexample	108
6.4.3. Good's Reply to Salmon	108
6.4.4. Humphreys' Theorem	109
6.5. Salmon On Positive Relevance	110
6.5.1. The Method of Successive Reconditionalization	110
6.5.2. Successive Reconditionalization and Indirect Causes	111
6.5.3. Problems With the Method of Successive Reconditionalization	112
6.6. Causes, Causal Factors, and Positive Relevance	113
6.6.1. Causes and Negative Causes	113
6.6.2. Causes That Are Not Causal Factors	114
6.7. Humphreys' Theory of Causality	115
6.7.1. Humphreys' Example	115
6.7.2. Critical Discussion of Humphreys' Theory	116
6.8. Causes and Causal Factors	118
7. SIMPSON'S PARADOX	120
7.1. Probabilistic Causality and Simpson's Paradox	120
7.2. Cartwright's Treatment of Simpson's Paradox	121
7.2.1. Informal Statement of Cartwright's Solution	121

7.2.2. A Formal Version of Cartwright's Solution	122
7.2.3. Modifications of Principle CC	124
7.2.4. Sufficient Causes and Necessary Causes	127
7.2.5. Philosophical Problems With Cartwright's Solution	128
7.3. Skyrms' Treatment of Simpson's Paradox	131
7.3.1. The Pareto-Dominance Condition	132
7.3.2. Cartwright's Rejection of the Pareto-Dominance Condition	132
7.3.3. Principle CC and the Pareto-Dominance Condition	133
7.3.4. Critical Discussion of the Pareto-Dominance Condition	133
7.4. Towards A Solution of Simpson's Paradox	136
8. SALMON'S THEORY OF CAUSALITY	138
8.1. Explication of Salmon's Theory	138
8.1.1. Causal Production and Causal Propagation	138
8.1.2. Causal Processes and Pseudo-Processes	139
8.1.3. The Transmission of Marks	139
8.1.4. The Principle of Causal Interaction	141
8.2. Critical Analysis of Salmon's Theory	142
8.2.1. The Counterfactual Aspect of Principle MT	142
8.2.2. Problems With Principle MT	143
8.2.3. Principle CI and Pseudo-Processes	147
8.2.4. Problems With Principle CI	148
8.2.5. The Relationship Between Principle MT and Principle CI	149
8.3. Different Types of Causal Interactions	151
8.4. Interactive Forks and Causal Screening Off	152
8.5. Salmon and Causal Connections	153
Appendix A. INTERPRETATIONS OF PROBABILITY	154
A.1. The Classical Interpretation of Probability	154
A.2. The Subjective Interpretation of Probability	156
A.3. The Personalist Interpretation of Probability	156
A.4. The Logical Interpretation of Probability	158
A.5. The Propensity Interpretation of Probability	161
A.6. The Frequency Interpretation of Probability	163

LIST OF FIGURES

Figure 3-1:	Causal Betweenness in a Causal Fork	38
Figure 3-2:	Causal Betweenness in a Causal Chain	38
Figure 3-3:	Marks in a Causal Fork	48
Figure 3-4:	Marks in a Causal Chain	49
Figure 3-5:	Causal Chain In Which Mark Appears	50
Figure 3-6:	Marking in an Interactive Fork	53
Figure 4-1:	Illustration of Example 2	70

ABSTRACT

Probability and Causality is a critical analysis of the problem of causality in indeterministic contexts. Most philosophers who have written about probabilistic causality feel that Hume's requirement of constant conjunction should be replaced by a requirement of positive statistical relevance. After arguing that a theory of probabilistic causality is necessary to account for many causal relations, Hume's theory of probabilistic causality is analyzed. Although Hume's theory is inadequate, it does form the basis for later discussions of probabilistic causality.

The first modern treatment of probabilistic causality is that of Hans Reichenbach, and it is discussed in detail since all later theories rely upon his basic intuitions. Reichenbach presented a proof that probabilistic definitions of causality were equivalent to the non-probabilistic analyses based on mark transmission. This proof is analyzed, and although it fails, several possible modifications of it are discussed. The next theory discussed is that of Patrick Suppes. It is shown that Suppes' theory is intrinsically defective, and that no minor modifications of his theory will be sufficient to solve the problems it faces. I. J. Good's quantitative theory of causation is then also shown to be defective.

Although almost all theories of probabilistic causality assume that causes raise the probability of their effects, there is no real defense of that requirement. The author attempts to clear up the confusion surrounding the discussions of this requirement by showing that two related but distinct causal concepts are being confused. The related problem of Simpson's paradox is then discussed, and it is shown that all proposed solutions to it face serious philosophical problems.

Salmon developed a theory of probabilistic causality which analyzes causal relations in terms of mark transmission instead of probability relations. The counterfactual aspect of mark transmission and causal interaction is closely examined.

Probability and Causality concludes with an appendix in which the various interpretations of probability are discussed in reference to developing a theory of probabilistic causality.

CHAPTER 1

PROBABILISTIC CAUSALITY

Since probabilistic causality is contrasted with what I shall call deterministic causality, we should begin by getting clear about what we mean by deterministic causality. By deterministic causality, I mean the conception of causality that is present in and advanced by theories such as Newtonian mechanics. The ideas of sufficiency and necessity are intimately connected with the idea of causation in Newtonian mechanics. This becomes clear if we consider Laplace's views on determinism. For Laplace, a cause is necessary and sufficient for its effect. This follows from the fact that Laplace claimed that if a being could know the present state of the universe and all laws of nature, that the being could also know the state of the universe at any time, whether past or future. The past or future state of the universe would be calculated using laws of nature, which are basically causal laws. The only way that this would be possible is if causes are necessary and sufficient for their effects. Hume, who was also influenced to a large degree by Newtonian mechanics, is also very clear that he believes that causes are necessary and sufficient for their effects. We often fail to recognize how instrumental Newtonian mechanics was in modifying our conception of causation, but after the development of Newtonian mechanics the idea that causes were necessary and sufficient for their effects was a very prevalent view.

1.1. The Problem of Identifying Deterministic Causes

Of course there were problems with this view, not the least being that for many events we could not identify a necessary and sufficient cause. In dealing with these situations, the general rule was to have faith that there was one, and to continue looking for it. The principle of sufficient reason was accepted, even though there did seem to be cases in which we were ignorant of the necessary and sufficient cause. Even when we thought we could identify the cause, because of limitations on describing it we seemed to be unable to show that it was necessary

and sufficient for the effect. This formed the basis of probability for Laplace. Laplace believed that probability is an expression of our ignorance. Even though every event has a necessary and sufficient cause, sometimes we are ignorant of it, and then we only have probabilities and not certainty. As more knowledge is accumulated we will discover the true causes of certain events, and then we will no longer need to rely on reasoning with probabilities. Any sort of causation that does not construe causes as necessary and sufficient for its effect is a false theory of causation, according to Laplace, that postulates "imaginary causes." These imaginary causes will gradually disappear as more knowledge is accumulated.

1.1.1. Hume's Need For Probabilistic Causes

Hume ran into serious difficulty when he was confronted with this problem. Being influenced by Newtonian mechanics, Hume had faith that every event had a necessary and sufficient cause. Hume believed this because of the enormous success that science was having. But Hume was also well aware that the vulgar did not believe this, and ascribed causality to all sorts of relations that fell short of necessity and sufficiency. This was a problem for Hume, because he claimed that any idea had its source in our experience, or our sense impressions. Thus he was forced to account for our idea of causation that is not deterministic in character. This is basically the problem of what to do when the connection between cause and effect is not one of constant conjunction. Most philosophers are well aware of Hume's analysis of deterministic causation, but very few are aware of his discussion of the causation of the vulgar, or probabilistic causation. Suffice it to say, Hume ran into serious difficulties when he tried to account for the idea of causation that is not deterministic. We shall discuss this in detail in the next chapter. For our present purposes, it is important to realize that Hume recognized that the deterministic causation of the philosophers was quite different from the causation of the ordinary person.

1.1.2. Mackie's Account of Causation

These problems are still present today. Modern philosophers are still trying to synthesize the idea that causation involves the ideas of sufficiency and necessity, with the fact that in many causal relations in the world we do not find either sufficiency or necessity. For example, a very popular modern account of causation is that of J. L. Mackie. Mackie (1974, p. 62) presents the idea of an INUS condition. An INUS condition is an *insufficient but non-redundant part of an unnecessary but sufficient condition*. This ploy attempts to account for the apparent fact that some causes are not necessary and sufficient for their effects, while still retaining the ideas of necessity and sufficiency. This way of looking at things sees a total cause as a string of disjunctions, each disjunct being a string of conjunctions. For example,

$$(A \& B \& C) \vee (D \& E \& F) \vee \dots \vee (X \& Y \& Z) \text{ causes } M,$$

where any of the members of the disjunctions are causes, or at least partial causes. For example, my drinking poison might be a cause of my dying, even though it is neither necessary for my dying, I could have a heart attack, nor sufficient for my dying, I could take an antidote. But according to Mackie's analysis, drinking the poison, not drinking an antidote, and no miracles, etc. is sufficient for my dying, and thus my drinking the poison is a partial cause of my dying. Without going into the details of this account, we should notice that it is built around the ideas of necessity and sufficiency: each partial cause is a necessary part of a sufficient condition. Thus this theory is to be considered a modification of the traditional deterministic conception of causation. Other philosophers talk about causal fields or *ceteras paribus* clauses. These are all attempts to fit the ideas of necessity and sufficiency with a real world that resists them. For the following discussion, by deterministic causation I mean to include any conception or analysis of causation which construes causation in terms of necessary and/or sufficient conditions.

1.2. The Problem of Ordinary Causal Concepts

1.2.1. Suppes' Examples

A similar tension that was always present in deterministic conceptions of causality is that our ordinary conception of causation is not deterministic in character: it seems as if deterministic causation is a scientific concept. We often apply the term "cause" to events that are neither sufficient nor necessary for their effects. This fact is brought out forcefully by Patrick Suppes. Suppes gives several examples that point out how our everyday way of speaking about causality is not deterministic in character. For example, when we say that someone's reckless driving is bound to lead to an accident, we are expressing a causal claim, but that claim in no way implies that reckless driving is either a necessary or sufficient cause of an accident. Other examples abound. Not dressing correctly in cold weather is said to cause one to catch a cold, but surely not everyone who does not dress correctly in cold weather catches a cold. In these, and other examples, Suppes notices that the cause tends to increase the probability of the effect occurring, even though it does not guarantee that the effect will occur. Thus Suppes claims that the core of causality lies not in necessity or sufficiency, but rather in an increase of probability.

1.2.2. A Response to Suppes

It appears that Suppes is correct in his claim that our everyday talk about causation is not deterministic, but from that it does not follow that people really believe that causes are not necessary and sufficient for their effects. Some might wish to claim that in cases when it appears as if causes are not sufficient or necessary, that we have not looked closely enough for the real cause. It might be claimed that people believe that there is a deterministic cause of every event, even if we can't find it. Ronald Giere (1979) argues this point in connection with certain medical examples. Even though we may speak of causation in cases in which the cause is not necessary or sufficient for the effect, we believe that if the situation were specified closely enough that all of the causes would be deterministic causes. To take Suppes' example, it might be claimed that the expression "his

"reckless driving is bound to lead to an accident" does not mean that we need some probabilistic notion of causation, but rather it means that it is likely that his reckless driving will be a deterministic cause of an accident. Thus one might reply to Suppes that people actually do believe that causes are deterministic, even if they use causal language loosely and speak of causes that are not deterministic. Probabilistic causality would be seen as dealing solely with imaginary causes, and not with the real structure of the world.

1.2.3. Suppes and Epistemic Causal Concepts

I suspect that Suppes would respond by claiming that whether or not most people believe that there are underlying deterministic causes is irrelevant. What is important is that people do speak of causation in situations in which the cause is not necessary or sufficient for the effect. Thus even if the world is ultimately deterministic, our ignorance of the actual causes results in our ascribing causality to situations in which there is no necessity or sufficiency. Suppes believes that causal relations are always relative to some theory or conceptual framework; causality is an epistemic notion that varies among conceptual frameworks. Causal relations do not exist in the world independently of our conceptual frameworks: thus it is irrelevant if upon some conceptual framework the world is ultimately deterministic. There are frameworks in which causality is not deterministic, and this must be accounted for.

1.3. Reichenbach and Indeterminism

Hans Reichenbach (1956) presents us with a different reason for needing a theory of probabilistic causality. Even before the development of modern quantum mechanics, he was claiming that probability, not the relations of sufficiency and necessity, played the major role in causality. He believed that even in classical physics there was a form of indeterminism, and that probabilistic causation was needed. The reasons for this were based upon a definition of determinism and the idea of testability. We won't discuss this in detail, except to note that Reichenbach believed that even Newtonian mechanics needed a conception of probabilistic causality.

1.4. Salmon and Probabilistic Causality

1.4.1. Salmon's Argument

Wesley Salmon (forthcoming) also claims that in classical Newtonian physics we are mistaken if we believe that causes must be deterministic. He presents several examples from physics in which causality is ascribed to events which are not related by sufficiency or necessity. Suppose we place an ice cube in warm water. We normally claim that placing the ice cube in warm water is a cause of the ice cube melting and the temperature of the water dropping. However, according to classical kinetic theory, there is a small probability that the ice cube will become colder, and that the water will become warmer. Thus placing the ice cube in warm water is not sufficient for it melting, and it is certainly not necessary. Thus even if determinism is true, there will be causes which are not necessary and/or sufficient for their effects according to Salmon.

1.4.2. A Reply to Salmon

One might reply to Salmon's example by emphasizing that it is based upon deterministic physics, and that if one knew the precise initial conditions, then one could predict with certainty whether the ice cube would melt in the warm water. The reason that placing the ice cube in warm water is not sufficient for its melting is because we have not specified all of the initial conditions that the ice cube and water are in; if we did specify these, placing the ice cube in warm water would be sufficient for its melting, or in rare cases it would be sufficient for its becoming colder. Examples such as this do not force us to adopt a theory of probabilistic causality, because it is possible in principle to account for the phenomenon with deterministic causality. It might be claimed that in claiming that placing ice cubes in warm water causes them to melt we are presupposing the appropriate initial conditions. Thus Salmon's example is not a conclusive reason to adopt a theory of probabilistic causality.

1.4.3. Salmon's Rejoinder

Salmon would not deny that many cases like the above can be handled by a very precise specification of the initial conditions. However, he would think that such a project is basically unnecessary, because the motivation for it comes from a strict adherence to determinism. Salmon would question the usefulness of building the concepts of necessity and sufficiency into our idea of causation, and he thinks that it is best to leave unnecessary requirements out of our analysis of causation. He would also point to several other examples from physics which are irreducibly statistical to support his point.

1.5. Modern Physics and Indeterminism

1.5.1. Causation and Indeterminism

I suspect most of us would still believe that all real causes are deterministic causes, if it were not for developments in modern physics. According to the theory of quantum mechanics, the world is not ultimately deterministic. Quantum mechanics tells us that there are some systems that are irreducibly statistical or probabilistic: no matter how well they are described, they cannot be made into deterministic systems. Thus if we hold that all causation is deterministic causation, we will have to claim that many of the events in these indeterministic systems are simply uncaused, or happen by chance.

Unfortunately for that view, physicists do speak of causality in indeterministic contexts. One can easily find examples in which physicists are studying the causal relations among events in indeterministic systems. It is obvious to physicists that it is coherent and legitimate to apply causal concepts in indeterministic settings, or in cases in which the cause is neither necessary nor sufficient for its effect.

Salmon also argues that it seems implausible to believe that all of our causal concepts will be meaningless if it turns out to be true that determinism is false. Whether or not determinism or indeterminism is true is an empirical question which seems to have no bearing on whether or not there are causal relations in the world. There are causal relations, independently of whether the world is deterministic or indeterministic, and thus it seems a mistake to build the theory of determinism into

that of causality as has been done in the past. All of this is support for the claim that the ideas of necessity and sufficiency are really not essential to the concept of causality.

1.5.2. Examples of Indeterministic Causation

Let us consider a few examples to see the role of causality in indeterministic systems. If an electron and a positron collide in an appropriate magnetic field, the result can be a variety of things: mu mesons, k^0 mesons, k^- mesons, or k^+ mesons may be produced. No matter how precisely the original interaction is specified, it is impossible to predict with certainty what type of particles will be produced. We can give certain probabilities for the different outcomes, but that is all. The production of these mesons is clearly a causal process, even though this is an indeterministic setting. Almost all of the interactions between subatomic particles follow this pattern. Consider hadron scattering reactions. A hadron scattering reaction is when two stable hadrons collide and two or more hadrons emerge. There are many possible outcomes of the collision; they are limited only by certain conservation laws. The results are indeterministic, in that no specific outcome can be predicted with certainty, even though probabilities can be assigned to the outcomes. So suppose two protons collide. The result could be the same two protons emerging, or two protons and other particles might emerge; we might have two protons and a neutral pion emerge. This is clearly a causal reaction, but it is also clearly indeterministic. Since deterministic theories of causation cannot handle these examples, a probabilistic theory of causation is needed.

As another example, suppose that we have an atom in an excited state, and that a photon impinges on it and causes the atom to decay to the ground state by emitting a photon. No one doubts that this is a legitimate case of causation, and indeed, this is just how some lasers work (Salmon forthcoming). However, the photon impinging on the atom is neither necessary nor sufficient for the atom to decay to the ground state. It is possible that the atom might decay to the ground state in the absence of the photon impinging on it, and it is possible that it might not decay to the ground state even though the photon impinges on it. According to quantum mechanics, this is an irreducibly statistical relation, and there is no way, even in principle, of describing the situation which will remove its probabilistic

indeterministic aspect. Since we speak of causation in situations such as these, we must be prepared to give an analysis of causality that will handle these situations.

1.6. Positive Relevance

These considerations show the need for a theory of causality that is different from the traditional theories which involve the ideas of sufficiency and necessity. Theories of probabilistic causality are an attempt to meet this need. Intuitively, most people think that a cause should raise the probability of its effect, and this is what most philosophers that write on probabilistic causality propose that we replace the requirement of sufficiency and necessity with. Positive statistical relevance is more important than sufficiency or necessity. Returning to Hume, the proposal is to replace the idea of constant conjunction with the idea of an increase in probability. The intuitive idea is that if A is a cause of B, then the occurrence of A should make B occur more often than it would have if A had not occurred. More precisely, we say that A is a cause of B if and only if the conditional probability of B on A is greater than the probability of B simpliciter: $P(B/A) > P(B)$. This is what many consider to be the essence of probabilistic causality.

A look at the examples we discussed earlier shows that positive statistical relevance, not sufficiency or necessity, was present in the examples we discussed. For example, reckless driving raises the probability of an accident, not dressing correctly raises the probability of sickness, and placing an ice cube in water raises the probability of its melting. There are also numerous medical examples to support this idea. If we look at almost any study in medicine, we find that the researchers are looking for relations of positive statistical relevance. When researchers try to decide if saccharin causes bladder cancer, they look for relations of positive statistical relevance, because the actual probability of getting bladder cancer, even if one takes large amounts of saccharin, is quite small (Giere 1979). But the presence of saccharin does significantly raise the probability of getting bladder cancer, so we say that it is a cause of bladder cancer. Similarly, one could look at the probability of getting lung cancer given that one smokes cigarettes heavily, as compared to the general probability of getting lung cancer. Certainly not everyone who smokes heavily gets lung cancer, but it has been shown that

smoking heavily does significantly raise the probability of getting lung cancer, and thus we claim that it is a cause of lung cancer. In almost all medical examples the researchers are dealing with positive statistical relevance, and not with relations of sufficiency and necessity.

Accordingly, many philosophers have attempted to analyze causation in terms of probability relations. Thus most theories of probabilistic causality are also probabilistic theories of causality, where by a probabilistic theory of causality I mean a theory that attempts to analyze causation in terms of probability relations. There are theories of probabilistic causality that do not attempt to analyze causality in terms of probability. In the coming chapters, we shall investigate both of these types of theories of causality in an attempt to discover the relationship between probability and causality.

CHAPTER 2

HISTORICAL ROOTS OF PROBABILISTIC CAUSALITY

2.1. Hume's Discussion of Causality

Just as all discussions of causality during the Middle ages took Aristotle as the starting point, so must all modern discussions of causality begin with the work of David Hume. Hume's masterful analysis of causality is so important that any serious discussion of causality must respond to his ideas. Thus in our study of probabilistic causality it will profit us to investigate his theory, and use it as a place to begin our discussion. Most people believe that Hume did not discuss probabilistic causality; we must ascertain the truth of this belief, and if it is true that he did not discuss probabilistic causality we must attempt to determine what he would have said about it had he been asked.

2.1.1. Hume's Theory of Deterministic Causality

Hume begins his discussion of causality by trying to discover where we get the idea of causation. This is important for Hume, because his strict empiricism led him to believe that all ideas are just representations of sensory impressions. Hume says that people generally think of causation as furnishing some sort of necessary connection between a cause and its effect. Thus he searches for the origin of this idea of necessary connection. He immediately notices that there is no logical connection between a cause and effect, or the idea of a cause, and the idea of its effect. One can easily conceive of the cause, and imagine some contrary of the effect to occur. But if there is no logical connection between a cause and effect, from where can the idea of a necessary connection arise? We can never observe any connection between the different events that are causally connected: all events seem to be separate from one another. Hume (1893, p. 58) notes that they are often conjoined together, but we never observe them connected. Since necessary connections in causal relations are not detected by the senses, he concludes that we

do not have any idea of a necessary connection in the objects themselves. All that he can discover in causal relations is that the cause and effect are always conjoined.

Although Hume does not believe that there is any necessary connection between a cause and effect, he still must find the source of our belief that there is such a connection. Hume says that when we first observe an event, we cannot determine what event will follow it; anything is possible. However, after we have seen one event follow another several times, we begin to believe that we can infer the existence of the later event from the occurrence of the earlier event, and to call the earlier event the cause of the later event. The observation of this constant conjunction between two events is the source of the idea of a necessary connection between a cause and effect. Although the idea of necessary connection could not arise from the observation of any single instance of the two events being conjoined, after several of them have been conjoined the mind acquires the idea of a necessary connection.

The reason that we cannot discover a necessary connection by observing two conjoined events is because the necessary connection is not in the objects, but rather it is in the mind. Hume (1893, p. 59) considered this to be one of his important discoveries. When the mind first observes two events conjoined, it has no tendency to infer the existence of the later from the existence of the earlier. However, after observing this constant conjunction several times, the mind acquires the habit to infer the existence of the effect from the existence of the cause. When the idea of the cause appears to the mind, the mind, because of habit, infers the existence of the effect. Thus the necessary connection between a cause and effect is not in the objects, but rather is only a habit of the mind that arises out of constant conjunction. This habit gives rise to a feeling that the cause and effect are connected, but this connection is only in our mind.

After making the discovery that the necessity in causal relations is in the mind, Hume proceeds to give definitions of a cause. Since the definitions in the *Treatise* are different than those in the *Enquiry*, I will state both versions of them. Let us begin with the first definition that Hume gives in the *Treatise* (1888, p. 170):

Definition 1: We may define a CAUSE to be 'An object precedent and contiguous to another, and where all the objects resembling the former are plac'd in like relations of precedence and contiguity to those objects, that resemble the latter.'

This definition defines the causal relation to mainly be based on the relationship of constant conjunction. It is also required that the cause precede the effect, and that the cause and effect be contiguous in space and time. Thus Hume cannot use the causal relationship to define temporal relationships in a causal theory of time.

Hume says that some people may find this first definition defective, because whether or not an event is a cause depends upon other events that are different from the cause in question. To remedy this, Hume (1888, p. 170) proposes a second definition:

Definition 2: A CAUSE is an object precedent and contiguous to another, and so united with it, that the idea of the one determines the mind to form the idea of the other, and the impression of the one to form a more lively idea of the other.

This definition contains Hume's version of the necessary connection present in causal relations.

The *Enquiry*, which was written after the *Treatise*, also contains two definitions of a cause. The first is:

Definition 3: ...we may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second (1893, p. 60).

The second definition given in the *Enquiry* (1893, p. 60) is as follows:

Definition 4: We may...form another definition of cause, and call it, an object followed by another, and whose appearance always conveys the thought to that other.

Immediately following the statement of the first definition in the *Enquiry*, Hume (1893, p. 60) gives another definition that he considers equivalent to definition 3:

Definition 5: [A cause is an object, followed by another] where, if the first object had not been, the second never had existed.

One immediately notices a difference between definition 5 and the previous two definitions. The main difference is that definition 5 is a counterfactual analysis, and definitions 3 and 4 are not. There is no analogue to definition 5 in the *Treatise*, and it is not mentioned by Hume as much as the other two are. Although this short definition of Hume's is the source of modern counterfactual analyses of causation, we shall for the most part be concerned with the other two

definitions of causation, because most writers on probabilistic causality have not given counterfactual definitions. For a modern counterfactual analysis of probabilistic causality see Fetzer and Nute (1979, 1980).

One interesting question that immediately arises concerns the differences between the definitions given in the *Treatise* and in the *Enquiry*. One striking difference is the inclusion of the requirement of spatial contiguity in the *Treatise*. At first it may appear as if temporal precedence is also left out of the definitions in the *Enquiry*, but it appears that the phrase "followed by" in definitions 3 and 4 imply temporal precedence of the cause. The absence of spatial contiguity in the *Enquiry* may have been due to Hume's desire to simplify the presentation of the ideas in the *Treatise*, or it may have been due to a change in Hume's earlier belief that it was necessary for causation. Hume does not give us the answer to this question. Otherwise, definitions 1 and 3 appear to be equivalent, although there may be a difference between definitions 2 and 4. Definition 4 simply requires that the idea of the effect follow the idea of the cause, but definition 2 also requires that the impression of the cause form a more lively idea of the effect. It is not entirely clear what it means to have a more lively idea of the effect. It might mean that the idea of the effect is more lively than it was before the idea of the cause was before the mind, or it might mean that the idea of the effect is more lively than the idea of the cause. It seems unlikely to me that Hume would require that the idea of the effect be more lively than that of the cause, so I will assume that he means that the idea of the effect must be more lively than it was before. Perhaps if the mind forms an idea that it did not have previously, the idea is more lively than it was. In this case it would seem that definitions 2 and 4 are equivalent except for spatial contiguity. The issue of whether an idea is more lively simply by being formed in the mind is an issue that we will return to when we discuss Hume's theory of probabilistic causality.

It is important to realize that for Hume, causation is actually necessary and sufficient causation: causes are necessary and sufficient for their effects. The definitions of a cause given ensure that all causes are sufficient causes. If a cause is always followed by its effect, then the cause is a sufficient cause. Although it is easy to see that Hume considered causes to be sufficient causes, it is more difficult to see that he also considered them necessary. The definitions of a cause given do

not seem to require that causes be necessary for their effects. However, immediately following his definitions of a cause, Hume (1888, p. 173–4) writes the following:

The same cause always produces the same effect, and the same effect never arises but from the same cause. This principle we derive from experience, and is the source of most of our philosophical reasonings.... There is another principle, which hangs on this, viz. that where several different objects produce the same effect, it must be by means of some quality, which we discover to be common amongst them. For as like effects imply like causes....

From these passages we can see that Hume did not a priori rule out necessary but insufficient causes, or sufficient but non-necessary causes: in Hume's experience all causes were necessary and sufficient for their effects, upon close inspection. It would not be very difficult to expand Hume's definitions of a cause to definitions of necessary but insufficient causes, or sufficient but non-necessary causes.

2.1.2. Hume's Theory of Probability

In order to discuss Hume's theory of probabilistic causality we must first look at Hume's theory of probability, and his ideas about indeterminism. In the *Treatise* Hume examines the various arguments for the thesis that every event must have a cause, and he finds them all inadequate. Thus Hume believes that we cannot prove that every event has a cause; the source of our belief that every event has a cause is our experience. One might object to the thesis that our experience informs us that every event has a cause, and claim that we commonly experience certain events that happen by chance. Hume (1888, p. 132) replies:

The vulgar, who take things according to their first appearance, attribute the uncertainty of events to such an uncertainty in the causes, as makes them often fail of their usual influence, tho' they meet with no obstacle nor impediment in their operation. But philosophers observing, that almost in every part of nature there is contain'd a vast variety of springs and principles, which are hid, by reason their minuteness or remoteness, find that 'tis at least possible the contrariety of events may not proceed from any contingency in the cause, but from the secret operation of contrary causes. This possibility is converted into certainty by farther observation, when they remark, that upon an exact scrutiny, a contrariety of effects always betrays a contrariety of causes, and proceeds from their mutual hindrance and opposition.... From the observation of several

parallel instances, philosophers form a maxim, that the connexion betwixt all causes and effects is equally necessary, and that its seeming uncertainty in some instances proceeds from the secret opposition of contrary causes.

This is a statement of belief in determinism. Since we can always find a necessary and sufficient cause for any apparently indeterminate event, Hume thinks that we have reason to believe that all events are determined. Although it may appear as if this thesis is inconsistent with his belief that we cannot prove that every event has a cause, upon closer inspection we find that the inconsistency is only apparent. Hume had shown that there was no logical reason to claim that every event had a cause, even though we may find by experience that it is true. The fact that every event has a cause is, for Hume, not an *a priori* fact about the world, but rather a contingent feature that it has.

Given Hume's belief in determinism, we must now consider his doctrine of probability and chance. Hume (1888, p. 171) thinks that chance is the absence of a cause:

'Tis the constant conjunction of objects, along with the determination of the mind, which constitutes a physical necessity: And the removal of these is the same thing with chance. As objects must either be conjoin'd or not, and as the mind must either be determin'd or not to pass from one object to another, 'tis impossible to admit of any medium betwixt chance and an absolute necessity.

In this passage Hume is claiming that there is causation, chance, and nothing else; there is no room left for probabilistic causality.

Since Hume believes that chance is the absence of a cause, and that all events have causes, we would expect him to also claim that there is no such thing as chance. However, he does not make this claim, and discusses reasoning about chances in some detail. Although he does not discuss this tension that arises from his doctrines, I suspect that the solution is to be found in his doctrine of causation. Hume was well aware that many people would not recognize certain causes in the world; that is why he spoke of secret causes. We might read Hume as presenting an epistemic conception of chance. Although the philosophers would realize that every event has a cause, many people would be unaware of them, and thus their minds would not make the transition from the cause to the effect. Thus there would be events due to chance, as far as these people are concerned. If my interpretation of Hume is correct, it is unfortunate that he did not differentiate

between chance as the absence of a cause, and chance as the ignorance of a cause. Once this distinction is made, Hume's ideas become much more coherent.

Hume then proceeds to endorse what is known as the classical theory of probability. The classical interpretation of probability defines probability to be the ratio of favorable to equally possible cases. We determine what the equally possible cases are by applying the principle of indifference, which says that two events are equally probable if the mind is indifferent between them, or has no reason to prefer one to the other. Hume (1888, p. 125) uses this principle when he says that "an entire indifference is essential to chance." Once we know the number of chances that are favorable and the number of chances that are possible, we can calculate the probability of a certain outcome. The classical interpretation of probability is discussed in further detail in the appendix.

Hume gives an account of our reasoning about probability in terms of his theory of belief. Suppose we are throwing a die that has one mark on four of its sides, and two marks on two of its sides. When we throw the die, we realize that it will land in such a way that one side will be facing up, although we will not know which side it will be. According to Hume (1888, pp. 124-30) our mind imagines the various possibilities, but since they are all equally probable, not one of them has any more force upon our mind than any of the others. However, several of the sides have the same number of marks on them, and they appear the same to the mind; hence the mind joins all of the images or ideas together which are similar in outcome. In our example, there are four ideas or images of a side with one mark on it facing up. These four ideas join together to form one idea or image that is stronger and more lively than they were individually. Similarly, the two ideas of a side with two marks on it facing up will join together to form an idea that is more lively and vivid than either of them were previously. At this point there are two images before the mind that differ in strength and which represent the different probabilities of the two outcomes. We see that in reasoning about probability the mind first reviews the various alternatives: each possible outcome has the same probability. Then the mind joins together the outcomes that are similar, to get different degrees of probability.

2.1.3. Hume's Discussion of Probabilistic Causation

Although Hume doesn't really believe that there is such a thing as probabilistic causality, he cannot deny that many people do seem to have the idea of probabilistic causation. Merely claiming that these people are ignorant of the secret causes does not solve the problem of how we have the idea of probabilistic causation. If Hume's empiricism is correct, he should be able to explain to us the source of this idea. In the *Treatise* (1888, pp. 124–30) he devotes several pages to a discussion of this problem, whereas in the *Enquiry* (1893, p. 47) he only spends about a page on it, and claims that the more we contemplate the problem, the more difficult it becomes.

According to Hume (1888, p. 132), both the deterministic causation of the philosopher and the probabilistic causation of the vulgar are founded on the same basic principles. Just as the idea of deterministic causation arose out of the relation of constant conjunction, the idea of probabilistic causation arises out of the relation of frequent conjunction. Hume claims that the idea of probabilistic causation or a "hesitating belief" about the future can arise in the vulgar in one of two possible ways. The first way it can arise is due to an *imperfect habit*. It is not clear what Hume meant by an imperfect habit. The first time we see two events conjoined, he noted, there would be little tendency to infer the one from the other. But after we have observed them conjoined many times we develop the habit. The strength of the habit will depend upon how many times we have observed them conjoined together. Similarly, the observation of frequent but not constant conjunction results in an imperfect habit. The strength of this imperfect habit will depend upon the frequency with which the events are conjoined. There are various ways to construe what Hume might have meant by an imperfect habit:

1. only some people's minds might make the transition from the cause to the effect,
2. our minds might make the transition only part of the time,
3. our minds might make the transition to a weaker belief.

The first possibility does not seem to be what Hume meant by an imperfect habit, because even where there is a perfect habit, the vulgar do not always make the transition. This interpretation would make all cases of causation into cases where

there is an imperfect habit, which Hume did not intend. Alternative 2 seems defective because upon this interpretation all habits would also be imperfect, since there would be a time when we did not make the transition, and a time, probably later, at which we do make the transition. This alternative would also imply that sometimes an imperfect habit gives rise to causal necessity, and sometimes it doesn't, since causal necessity consists in the feeling of necessity. Thus the third alternative seems to be the best choice as to what Hume meant by an imperfect habit. When the mind has an imperfect habit to go from one idea to another, the latter idea will not be as lively and vivid as it would if it were a perfect habit. Thus we see that the first way we can reason probabilistically is to have an imperfect habit that gives us a "hesitating belief" about the future.

However Hume (1888, p. 133) thinks that this is not the way that the inferences concerning probabilistic causality usually take place. When we make an inference from habit, we do not delay and contemplate the various alternative outcomes; the transition of the mind is made quickly without reflection. However Hume thinks that in inferences concerning probabilistic causality there is usually some sort of deliberation. Thus the second way in which we can have a "hesitating belief" about the future is not based directly on habit, but rather it arises indirectly from habit.

Hume first notices that we have a habit to transfer what has happened in the past into the future. This is a habit, and is not based on any inference. Thus when we are contemplating what will happen in a certain situation, the mind will tend to transfer what has happened in the past into the future. If in the past there have been several contrary effects from a cause, the habit to transfer the past to the future presents the mind, not with the idea of one object or event, but rather with a number of different ideas and images. The different ideas and images that are presented to the mind are all of the different possibilities that might happen.

So first the mind separates, or diverges from the image of the cause into the images of the various possible effects. The mind then combines various images together. This is a result of Hume's theory of probability, which holds that an image of each "chance" is presented to the mind. Thus images of some of the possible effects are presented more than once, and the really probable effects will

most likely have several images of them. The images that are similar to one another are then combined into one stronger image. Hume (1888, p. 134) gives an example in which he has observed that only nineteen out of twenty ships return to port after setting out to sea. If he then observes twenty ships setting out to sea, his mind will form the idea of nineteen ships returning safely and one as perishing. The reason the mind does this is that it has a habit to transfer the past to the future, which includes transferring past frequencies to the future also.

However, we will seldom see exactly twenty ships setting out to sea, and instead we will usually observe a single ship leaving port, and will wonder if it will return safely. The problem that arises in this situation is analogous to the problem of the single case in probability theory. When the single ship sets out to sea, Hume believes that all of the possible effects are presented to the mind. In this case it would seem that there are twenty equally probable effects or chances. Thus the mind would form twenty images: nineteen of a ship returning safely, and one of a ship perishing. Since the nineteen images agree with one another, they unite together and form a more lively and strong idea of a ship returning safely. Hume (1888, p. 137) says:

The only manner then, in which the superior number of similar component parts in the one can exert its influence, and prevail above the inferior in the other, is by producing a stronger and more lively view of its object. Each part presents a particular view; and all these views uniting together produce one general view, which is fuller and more distinct by the greater number of causes or principles, from which it is deriv'd.

We thus see that when the idea of a probabilistic cause appears to the mind, the mind first has images of all of the equally possible effects. The mind then combines the similar ideas together to make some of the ideas more lively and vivid than others.

It is interesting that Hume does not say what faculty of the mind brings all of the various equally possible effects before the mind. The fact that they do appear before the mind is certainly due to habit, but it appears as if more than just habit is involved. It seems as if memory would also have to be involved, in order to remember past frequencies. Although Hume does not mention memory, it appears as if it is operable here.

One interesting problem concerns the mind's ability to bring forth all of the

equally possible effects given the idea of the cause. In cases such as throwing a die, it seems obvious what the equally probable effects will be: each side of the die facing up is equally probable. However, in the case of nineteen out of twenty ships returning safely to port, it is not obvious that there were twenty equally possible effects. It seems as if the only way that we can know the number of equally probable outcomes is by experience with frequencies. Thus it seems that the real basis for Hume's theory of probability, in which equally probable "chances" appear to the mind, is the memory of past frequencies which will tell us what the equally probable "chances" are.

We see then that most of our inferences concerning probabilistic causes are based on habit. There is deliberation in the sense that after the mind moves to the ideas of the various possible effects, by habit, the mind joins all of the similar ideas together. The beliefs that the mind finally arrives at are uncertain: they are not as lively as they would have been had the cause been a deterministic cause. We should also notice that Hume is clear that the mind is able to distinguish between different probabilities, even if they differ only minutely. From the preceding discussion we can see that contrary to common opinion, Hume did have a theory of probabilistic causality.

2.1.4. Critical Discussion of Hume's Theory of Probabilistic Causation

We have previously discussed Hume's three definitions of causation, and his analysis of probabilistic causation. It will now profit us to consider how Hume's analysis of probabilistic causation compares with his definitions of necessary and sufficient causation. We immediately notice that Hume's discussion of probabilistic causality leaves out all reference to spatial contiguity. Thus his analysis will be closer to the definitions given in the *Enquiry* than to those given in the *Treatise*.

One striking feature of Hume's analysis of probabilistic causation is that it contains psychological elements. The analysis depends upon the mind forming several images, and then regrouping them. Thus we are tempted to say that Hume's analysis of probabilistic causality is much closer to definitions 2 and 4, than to his first definitions, 1 and 3. We also notice that Hume's first account of probabilistic causation is closer to definitions 2 and 4 than his second account is. Hume's first account was that the mind has an imperfect habit to move from the cause to a

weaker belief of the effect. This appears to be closer in form to the definitions of deterministic causality than his more complicated analysis that is only based on habit indirectly does.

Hume's second analysis relies on habit differently than his first analysis does. Habit determines the mind to form all of the various equally possible possibilities before it. Definitions 2 and 4 tell us that the mind is determined to form the idea of the effect that always follows the cause; Hume's second analysis of probabilistic causality claims that the mind is determined to form the ideas of all of the effects that might possibly occur, or have occurred in the past. Thus we see that this too is a weakening of the analysis of deterministic causation, but here the habit is not weakened, but rather what the mind is determined to is weakened, or diversified. The habit is just as strong: the mind transfers past frequencies into the future. Yet now the weakening is in the form of what is transferred into the future, and not in the transferring itself. So Hume's second analysis of probabilistic causality is also very similar to his analysis of deterministic causation.

We should notice that whereas Hume offers *definitions* of deterministic causes, he does not actually give us a *definition* of a probabilistic cause. It seems that since Hume believed that all cases of probabilistic causality were really cases of secret or hidden deterministic causation, that he did not even attempt to give a definition of a probabilistic cause. Hume never used the term "probabilistic cause," and it appears as if he was not very concerned about defining when something was a probabilistic cause. Instead, Hume used the phrase "reasoning from contrariety," and it appears as if he was trying to account for our reasoning about cases where contrary events have followed from a single cause. Basically Hume has just described the psychological processes that occur when we judge something to be a probabilistic cause. In a nutshell, Hume has said that we project past frequencies into the future, and given the idea of a probabilistic cause, we will have ideas of all of the possible effects in proportion to their probability. But this is not a definition of probabilistic causation. If we expand that into a definition of a probabilistic cause we get something like the following:

Definition 6: A is a probabilistic cause of B if, whenever the mind has the idea of A, the idea of B will be one of the ideas that are formed in the process of transferring the past into the future.

This seems to be the analysis of probabilistic causality that we get from Hume's second discussion of the reasoning from contrariety.

However it is fairly easy to see that definition 6 is, given the rest of Hume's theory, simply equivalent to the following:

Definition 7: A is a probabilistic cause of B if at sometime in the past B has been observed to follow A.

To see that these are equivalent, consider the following: If A is a probabilistic cause of B according to 7, then at some time in the past B has been observed to follow A. But if B has been observed to follow A then the mind will form the idea of B when it has the idea of A. The mind does this because given the idea of a cause, it will form the idea of all of the possible effects, which are the things that have followed the cause in the past. Thus if A is a probabilistic cause of B according to definition 7, it is also a probabilistic cause of B according to definition 6. This means that when the mind has the idea of A, the idea of B will be one of the ideas that is formed as a possible effect in the process of transferring the past to the future. The reason the mind forms the idea of B as a possible effect is because B has followed A in the past. Thus we see that if A is a probabilistic cause according to definition 6, then it is also a probabilistic cause according to definition 7. Hence the two definitions are equivalent.

We could also expand Hume's first analysis of reasoning about probabilistic causes into a definition of probabilistic causation:

Definition 8: A is a probabilistic cause of B if, whenever the mind has the idea of A, it has an imperfect habit to form the idea of B.

Both definitions 8 and 6 differ from definitions 2 and 4 in that definitions 2 and 4 both require that both the cause and effect *occur*. This requirement is not built into definitions 6 and 8. We could easily add that requirement if we so desired. We could also add the requirement of spatial contiguity if desired.

However, definitions 6, 7, and 8 are all deficient. The only reason I will discuss here is that they do not distinguish between positive probabilistic causes and negative probabilistic causes. Upon these definitions, if A is a probabilistic cause of B, then A is also a probabilistic cause of not-B. But this result seems unintuitive. If we say that driving while intoxicated is a probabilistic cause of accidents, we really don't want to also say that it is a probabilistic cause of not having accidents.

Similarly, we don't want to say both that smoking is a probabilistic cause of lung cancer, and that smoking is a probabilistic cause of not getting cancer. We have an intuitive idea that a cause somehow helps to bring about the occurrence of its effect; the cause is partly responsible for the effect occurring. But it is easy to see that if A helps to bring about B, then it is not the case that A helps to bring about non-B. We might argue that if A helps to bring about B, then A cannot help to bring about not-B, because in helping to bring about B, A hinders the bringing about of not-B. Since these intuitions seem to be part of our concept of causality, we must claim that the above formulations of Hume's theory are inadequate.

We can salvage the above formulations of Hume's discussion of probabilistic causality by claiming that Hume was not giving definitions, which are necessary and sufficient conditions. Instead we could view the above formulations as giving necessary conditions of probabilistic causality. Upon this construal, the above formulations appear much more reasonable.

Perhaps we can modify definitions 6 and 8 to account for the previous difficulties. To do this we should inquire into what we normally mean when we say that A is a probabilistic cause of B. We have already seen that part of what is meant is that A helps bring about the occurrence of B. Another part of what may be meant is that A makes it more likely that B will occur. In our previous example, we would say that driving while intoxicated makes it more likely that one will have an accident. It seems plausible that the likelihood of an event occurring is just the probability of the event occurring. Since Hume believes that an increase in probability is just an increase in the liveliness and vividness of the idea, we could give the following definition of a probabilistic cause:

Definition 9: A is a probabilistic cause of B if and only if the idea of A is always followed by the idea of B, and the idea of B is always stronger than it was before the idea of A occurred to the mind.

This definition is similar to both 6 and 8; like definitions 6 and 8, the idea of B always follows the idea of A. However, definition 9 differs from definitions 6 and 8 in that it requires that the idea of the effect be stronger after the idea of the cause is presented to it. This, in effect, says that the cause raises the probability of the effect.

However a puzzle arises when we compare definitions 6, 8, and 9 with 2 and

4. It looks as if definition 9 is equivalent to definition 4, which would indicate that something is wrong with the definitions we have presented. First we notice that a deterministic cause is also a probabilistic cause. If A is a deterministic cause of B, according to definition 4, then the mind always forms the idea of B after having the idea of A. This alone would make A a probabilistic cause of B according to definition 6. Earlier we discussed whether definitions 4 and 2 were equivalent, except for the requirement of spatial contiguity. If we believe that they are equivalent, then we also believe that a deterministic cause makes the idea of the effect more lively: this follows from definition 2. Thus A would also be a probabilistic cause of B according to definition 9. A would not be a probabilistic cause of B according to definition 8, because 8 requires that the habit of the mind be an imperfect habit, which neither Hume nor I clearly explained. If we require that an imperfect habit not be a perfect habit, then A would not be a probabilistic cause of B according to 8. The definitions of habits and imperfect habits need not concern us now because Hume himself thought that this was a rare form of probabilistic reasoning. What is important for us to notice is that if A is a deterministic cause of B according to definition 2 or 4, then it is also a probabilistic cause of B according to definitions 6 and 9. This should not surprise us, because deterministic causation may just be a limiting case of probabilistic causation.

The problem is that if A is a probabilistic cause of B according to definition 6 or 9, then it is also a deterministic cause of B according to definitions 2 and 4. Suppose that A is a probabilistic cause of B according to definition 6 or 9. Then the idea of A is always followed by the idea of B. This happens because the mind remembers that B has followed A in the past, and it forms ideas of all possible effects when it has the idea of the cause. We also say that the mind is determined to do this, because it has a habit to transfer the past to the future. Thus every time the idea of A occurs, the mind forms the idea of B. But this is just the requirement of definition 4. We can also conclude that A would be a cause of B according to definition 2, because the appearance of A would make the idea of B more lively. Thus we see that anything that is a probabilistic cause according to definitions 6 or 9 is also a deterministic cause according to definitions 2 and 4.

This result tells us that either something is seriously wrong with Hume's

definitions of deterministic causes, or something is wrong with our formulation of probabilistic causes; certainly not all probabilistic causes are also deterministic causes. We can easily see that it is Hume's definitions 2 and 4 that are defective. Even if we don't rely upon my formulation of definitions 6 and 9 as a reconstruction of Hume's theory of probabilistic causality, we can see that 2 and 4 are defective. Consider Hume's discussion of our reasoning about probabilistic causality. When a probabilistic cause occurs, the mind forms ideas of all the possible effects. The mind always does this when presented with the idea of that probabilistic cause. If we take any of those possible effects, we find that the idea of it always follows the idea of the cause. But the intuition that 2 and 4 were built around is that whenever one idea always follows another, the earlier is the idea of the cause, and the latter is the idea of the effect. But Hume's own discussion of our reasoning about probabilistic causality shows that intuition or analysis to be false. It is a surprising conclusion to find that for Hume, who didn't really believe in probabilistic causality, that all probabilistic causes would also be deterministic causes.

We might attempt to modify definition 4 to account for this problem:

Definition 10: A cause is an object, followed by another, and whose appearance always conveys the thought to the other, and never to any contrary of the other.

This definition is the same as definition 4 except that it is required that the mind never form the ideas of contrary effects, which it would do if the cause were probabilistic. The definition of deterministic causality in the *Treatise*, definition 2, could be modified similarly. With the modification, no probabilistic causes will also be considered deterministic causes, although deterministic causes will still be probabilistic causes.

However definition 10 cannot be a satisfactory definition of deterministic causation, because of the following example due to Salmon. Suppose that just before I leap out of a 7th story window in an office building, I have a fantasy about floating across the street to the 7th floor of another building. This idea is contrary to the idea of falling to the ground, yet we would want to claim that jumping out of the window is a sufficient cause of my falling to the ground with disastrous consequences. Thus definition 10 would exclude many straightforward cases of deterministic causation from being causal relations. It seems to be possible

to fantasize about contrary outcomes when observing a sufficient cause; because of this, definition 10 is inadequate. A correct definition of deterministic causation consistent with the rest of Hume's philosophy is difficult to formulate. We will not investigate this any further now, except to note the difficulty.

The above definitions of probabilistic causality may be viewed as weakening definitions 2 and 4. We should also investigate whether a satisfactory definition of probabilistic causality can be given by weakening definitions 1 and 3. Although Hume did not give any analysis of probabilistic causality similar to 1 and 3 as he did for 2 and 4, we can attempt to discover what he might have said.

The essence of definitions 1 and 3 is that A is a cause of B if all objects that resemble A are followed by objects that resemble B. It is obvious that the "all" must be weakened. One obvious weakening of "all" would be to replace it with "most." The claim would then be that A is a probabilistic cause of B if *most* objects similar to A are followed by objects similar to B. However, even if this looks to be the most plausible weakening of definitions 1 and 3, it has serious problems with low probability events. We may consider untreated syphilis to be a probabilistic cause of paresis, but only a very few cases of untreated syphilis will ever result in paresis. It is simply false that most cases of untreated syphilis are followed by paresis, even though untreated syphilis is a probabilistic cause of paresis. It seems to me that the intuition behind this definition of probabilistic causality is that one should be able to infer the existence of the effect from the existence of the cause. Hume assumed that the necessary connection in causation allowed this inference, and in cases of deterministic causation it is possible to infer the existence of the effect from the existence of the cause. However, the ability to make this inference does not seem to be built into our concept of causation. It certainly seems like there are genuine cases of probabilistic causation which do not allow the effect to be inferred from the cause because of low probability relations. Thus this proposed weakening of definitions 1 and 3 is seen to be defective.

Earlier we decided that a crucial feature of probabilistic causality is that the cause raises the probability of the effect. Perhaps we could just weaken definitions 1 and 3 to require this feature. But this raises the question of whether the classical theory of probability is adequate for this purpose, which is discussed in the appendix. Hume needs a theory of probability that does not rely upon secret

powers, and is empirically testable. With such a theory of probability it does seem as if he could claim that a probabilistic cause is a cause that raises the probability of its effect. It does appear as if a probabilistic cause increases the propensity of the effect to occur. A theory of probabilistic causality seems to be intimately connected with the idea that the strength of a causal disposition is increased. It would then appear that definition 9 would be based upon this definition in much the same way that definition 2 appears to be based on definition 1 and definition 4 appears to be based on definition 3.

We saw that definition 5, which involved counterfactuals, was much different than the other definitions of deterministic causation. We might also try to give a counterfactual definition of probabilistic causation. The essence of definition 5 is that if the cause did not occur, then the effect would *never* have existed. The obvious weakening would be the "never." We cannot weaken it to say that if the cause did not occur, then the effect would probably not occur without running into the same problem with low probability events that we just discussed. It seems as if the proper counterfactual definition would be something like, if the cause did not occur, then the probability of the effect would be lower. Although we should note that propensities and probabilities are often claimed to have counterfactual elements, we will not devote much effort to discussing this attempt at a counterfactual definition, because the counterfactual definition of causation is not really a main part of Hume's analysis of causality.

2.1.5. Concluding Remarks on Hume

Now that we have developed a Humean theory of probabilistic causation, we must inquire into its shortcomings. One major shortcoming is Hume's failure to distinguish between genuine causes and spurious causes. This shortcoming was inherited from his theory of deterministic causality. Very early in the *Treatise* (1888, p. 4), Hume is arguing that if two events are constantly conjoined, one is the cause of the other. He develops this in more detail later, where the constant conjunction is offered as proof that one is the cause of the other, and temporal priority determines which one is the cause. However, this principle totally ignores the possibility that the connection may be due to a common cause. This difficulty is carried over into the analysis of probabilistic causality. We must remember that

frequent conjunction can be the result of a common cause just as easily as constant conjunction can..

Other problems that face Hume's account include that of distinguishing between positive and negative causes. Hume did not discuss this problem at all, and it certainly needs to be discussed. Another very serious problem is the fact that Hume's discussion of probabilistic causality depends heavily upon his theory of belief. It would be much more satisfying to have a theory that did not depend upon some psychological theory. Our attempt at weakening definitions 1 and 3 was aimed at a non-psychological definition of probabilistic causality, but this analysis is still in its early stages. More work would have to be done on the theory of probability to see if this analysis is plausible.

So far I have attempted to both explore and develop Hume's reasoning about causes that appear to be probabilistic in character. Although it is easy to see the shortcomings in Hume's discussion, we must not lose sight of the insight it has given us. Hume's ideas are intimately tied up with any analysis of causation, and his ideas about probabilistic causality have been neglected by most writers. But any theory of probabilistic causality must face the difficulties faced by Hume. We shall find that later writers will devote considerable effort into avoiding the problems that Hume faced in reasoning about probability and causation. In discussing the various theories of probabilistic causality it will be wise to remember what Hume (1893, p. 47) said at the end of his discussion of it in the *Enquiry*:

Let any one try to account for this operation of the mind upon any of the received systems of philosophy, and he will be sensible of the difficulty. For my part, I shall think it sufficient, if the present hints excite the curiosity of philosophers, and make them sensible how defective all common theories are in treating of such curious and sublime subjects.

2.2. Mill's Theory of Causality

After Hume, the next philosopher to have anything significant to say about causation was John Stuart Mill. In his *A System of Logic*, Mill discusses causation in detail. Like Hume, Mill believed that causes were sufficient for their effects; it is not clear whether he also considered them necessary as well. Since Mill thought that all causes were sufficient for their effects, he must deal with the problem of why it often appears as if causes are not sufficient for their effects.

2.2.1. Mill's Defense of Deterministic Causality

Mill attempts to solve this problem by attempting to fill in the relevant facts in such a way that the cause is sufficient for the effect. In defense of the invariable connection between cause and effect, Mill (1916, p. 214) writes:

It is seldom, if ever, between a consequent and a single antecedent that this invariable sequence subsists. It is usually between a consequent and the sum of several antecedents; the concurrence of all of them being requisite to produce, that is, to be certain of being followed by, the consequent. In such cases it is very common to single out one only of the antecedents under the denomination of Cause, calling the others merely Conditions. Thus, if a person eats of a particular dish, and dies in consequence, that is, would not have died if he had not eaten of it, people would be apt to say that eating of that dish was the cause of his death. There needs not, however, be any invariable connection between eating of the dish and death; but there certainly is, among the circumstances which took place, some combination or other on which death is invariably consequent: as for instance, the act of eating of the dish, combined with a particular bodily constitution, a particular state of present health, and perhaps even a certain state of the atmosphere; the whole of which circumstances perhaps constituted in this particular case the conditions of the phenomenon, or, in other words, the set of antecedents which determined it, and but for which it would not have happened. The real Cause is the whole of these antecedents; and we have, philosophically speaking, no right to give the name of cause to one of them exclusively of the others.

According to Mill the reason that causes may appear to not be sufficient for their effects is because we are only looking at a partial cause instead of the total or complete cause. When we look at the complete cause, according to Mill, the cause is sufficient for its effect. Mill is suggesting that we reserve the term "cause" for complete causes, which will have the result that all causes are sufficient for their effects. This has the result that causes are indefeasable: a cause cannot be defeated by the presence or absence of other factors. By bringing all of these other factors into the cause, Mill has explained how causes can be sufficient for their effects, and yet not appear to be so.

Mill claims that we know that every event has a sufficient cause by using induction by enumeration. It is not an a priori truth that every event has a sufficient cause, but we have found that every event we have examined closely does. This is basically the same justification that Hume used to justify the law of universal causation.

Although at the time of Hume and Mill it may have appeared as if every event has a sufficient cause, that belief is not as widely accepted now. Since the development of quantum mechanics it seems reasonable to believe that the world is ultimately indeterministic in nature, which means that not every event has a sufficient cause. We now have reason to believe that not every event we examine closely will have a sufficient cause. According to Mill, this is equivalent to claiming that some events are uncaused, since all causes are sufficient causes. Mill did not have the idea of a probabilistic cause, and thus if an event lacks a sufficient cause, it is an uncaused event.

2.2.2. Mill's Methods and Deterministic Causality

When one examines the methods Mill presented to determine causal relations, one finds that they are intimately connected with the idea of sufficient causation. Since Mill did not consider probabilistic causality to be true causality, the inductive methods that Mill developed do not pick out probabilistic causes. Mill was aware that his methods for finding causal relations presuppose that every event has a sufficient cause. Mill (1916, p. 369) writes:

The validity of all the Inductive Methods depends on the assumption that every event, or the beginning of every phenomenon, must have some cause, some antecedent, on the existence of which it is invariably and unconditionally consequent. In the Method of Agreement this is obvious; that method avowedly proceeding on the supposition that we have found the true cause as soon as we have negatived every other. The assertion is equally true of the Method of Difference. That method authorises us to infer a general law from two instances; one, in which A exists together with a multitude of other circumstances, and B follows; another, in which A being removed, and all other circumstances remaining the same, B is prevented. What, however, does this prove? It proves that B, in the particular instance, cannot have had any other cause than A; but to conclude from this that A was the cause, or that A will on other occasions be followed by B, is only allowable on the assumption that B must have some cause; that among its antecedents in any single instance in which it occurs, there must be one which has the capacity of producing it at other times. This being admitted, it is seen that in the case in question that antecedent can be no other than A; but, that if it be no other than A it must be A, is not proved, by these instances at least, but taken for granted. There is no need to spend time in proving that the same thing is true of the other Inductive Methods. The universality of the law of causation is assumed in them all.

Mill is claiming that the method of difference, and the other methods, assume that every event has a cause. If this is not assumed, we can never infer that a certain event is the cause of another, although we can infer that a certain event is the cause if there is a cause at all. Mill's reasoning here depends upon all causes being sufficient causes. If we concede that some causes are probabilistic and not sufficient, then it does not follow that we can infer what the cause is, if there is a cause at all. If not all causes are sufficient causes, then the method of difference still breaks down, because one of the other factors other than A may be the cause, even though the effect is present when A is present and absent when A is absent. Thus we see that Mill's methods presuppose the idea of sufficient causation.

2.2.3. Modification of Mill's Methods

One striking feature of Mill's methods is that they are seldom used in the manner in which he described. This is because it is not often that we can find examples of causation in which the effect always follows the cause, and never follows the absence of the cause. If we look at an experimental science which is trying to determine causal relations, we notice that few causal relations fit this pattern. For example, if one looks at current research being done in medicine, one finds that few, if any, causes are necessary and sufficient for their effects. Thus unless we reinterpret Mill's methods, we will be forced to conclude that they are inapplicable to most scientific research being done today. This fact is brought out in detail by Salmon (1981).

Salmon notes that Mill's methods can be modified slightly to account for these cases of non-deterministic causation. What we notice in cases such as these is that, although the effect does not always follow the cause, the effect follows the cause more often than the effect follows the absence of the cause. If C is the cause and E is the effect, then we notice that E follows C more often than E follows $\neg C$. Of course, if we reinterpret Mill's methods in this manner, they are no longer applicable to just two experiments; many situations will have to be observed to determine if the effect really follows the cause more often than it follows the absence of the cause. But this result seems to be in accord with accepted scientific practice. Very seldom, if ever, is a conclusion made on the basis of two instances, one in which the cause is present, and one in which it is not. Thus we might view this as a necessary addition to Mill's theory.

This modification of Mill's methods is necessary in order to enable them to be of any use in determining causal relations. Mill was preoccupied with deterministic causality, and thus his methods were tailored to deterministic causation. Since it is no longer accepted that all causes are deterministic, some modification is needed to account for actual scientific practice. From this modification of Mill's methods, it is only a small step to using the frequencies to infer probability values, and using probability relations to determine causal relations. It is to the credit of Mill that his inductive methods are able to be modified to account for probabilistic causation. Mill's methods are far from being a complete and satisfactory account of probabilistic causation, but we must not neglect the insight that they give us. A more detailed treatment of Mill's methods and the modifications necessary to make them applicable to scientific research can be found in Salmon (1981).

2.3. Gasking's Treatment of Causality

2.3.1. Brief Statement of Gasking's Theory

A modern theory of causality that has some relation to the ideas of Mill is that of Douglas Gasking. Gasking believes that the fundamental idea behind causation is that of producing and manipulation. Gasking (1972, p. 522) presents two points about causation:

First: that one says "A causes B" in cases where one could produce an event or state of the A sort as a means to producing one of the B sort. I have, that is, explained the "cause-effect" relation in terms of the "producing-by-means-of" relation.

Second: I have tried to give a general account of the producing-by-means-of relation itself: what it is to produce B by producing A. We learn by experience that whenever in certain conditions we manipulate objects in a certain way a certain change, A, occurs. Performing this manipulation is then called: "producing A". We learn also that in certain special cases or when certain additional conditions are present, the manipulation in question also results in another sort of change, B. In these cases the manipulation is also called "producing B", and, since it is in general the manipulation of producing A, in this case it is called "producing B by producing A".

The core of Gasking's theory is that A is a cause of B if by manipulating the

occurrence of A we manipulate the occurrence of B. It should be obvious that this account presupposes a deterministic conception of causation.

2.3.2. Modification of Gasking's Theory

The similarity of Gasking's proposal with Mill's method of difference suggests that we can modify Gasking's theory to account for probabilistic causality in much the same way that Mill's methods were modified. In cases in which A is a probabilistic cause of B, manipulating the occurrence of A will not always manipulate the occurrence of B. However, it seems reasonable to think that manipulating the occurrence of A will manipulate how often B occurs. Thus we can modify Gasking's theory to account for probabilistic causality by requiring that we manipulate how often the effect occurs by manipulating how often the cause occurs.

This modification of Gasking's theory is the natural way to account for probabilistic causality within his theory. Like Mill's methods, it appears as if Gasking's proposal is used extensively by working scientists; several engineers that I know are convinced that Gasking's theory captures the essence of causality. But in spite of its practical appeal, it is not an adequate account of causality. There are many cases of causation in which we cannot manipulate the cause, or the effect. Thus Gasking's theory cannot account for probabilistic causality as it stands, even though it does capture some of the intuitive ideas behind our causal notions.

CHAPTER 3

REICHENBACH'S TREATMENT OF CAUSALITY

The first modern treatment of probabilistic causality was Hans Reichenbach's *The Direction of Time*, which appeared posthumously in 1956. In 1925 Reichenbach (1978, pp. 81–119) did publish an article, "The Causal Structure of the World and the Difference Between Past and Future," dealing with probability and causality. Although this should be considered a modern discussion of the problem, Reichenbach considered his later discussion to be superior to the earlier, and we will limit our discussion to his later works. In *The Direction of Time* Reichenbach gave various different methods of constructing a causal net, which played a role in determining the direction of time. Reichenbach first used probability relations to define a relation of causal betweenness. He then developed probabilistic definitions of conjunctive forks and causal relevance. We shall begin our discussion of Reichenbach by looking at his discussion of causal betweenness; we will then consider the actual definitions of probabilistic causality that he proposed.

3.1. The Relation of Causal Betweenness

3.1.1. Definition of Causal Betweenness

The relation "causally between" is supposed to capture the structure of the causal net; it is supposed to tell us when two events are causally connected in a certain way. When we say that B is causally between A and C, we are saying that there are causal processes that connect A and B and B and C. Reichenbach gives two examples: a causal chain, and a causal fork. Both of these arrangements exemplify the relation of causal betweenness, although in different ways.

Reichenbach's (1956, p. 190) definition of the relation "causally between" consists of three requirements:

Definition 1: An event A_2 is causally between the events A_1 and A_3 if the relations hold:

$$1 > P(A_3/A_2) > P(A_3/A_1) > P(A_3) > 0, \quad (3.1)$$

$$1 > P(A_1/A_2) > P(A_1/A_3) > P(A_1), \quad (3.2)$$

$$P(A_3/A_1 \& A_2) = P(A_3/A_2). \quad (3.3)$$

We will symbolize this relation by $\text{btw}(A_1, A_2, A_3)$. In the above equations, and in the rest of our discussion, I will rewrite Reichenbach's probability notation in standard form. Thus $P(A, B)$ will be written as $P(B/A)$.

3.1.2. Explication of Definition 1

3.1.2.1. Positive Relevance. This definition formalizes the principle that the closer one gets in a causal chain to an effect, the better one is able to predict the occurrence of the effect. Equations (3.1) and (3.2) claim that regardless of the direction from which we approach an event, the closer we get to the event, the higher its probability becomes. Let us consider a simple example which illustrates what Reichenbach was trying to get at with this definition of causal betweenness. Suppose that we have three events which are causally connected in such a way that A_1 causes A_2 , and A_2 causes A_3 . Now suppose that we know that A_1 has occurred; then we can predict with probability $P(A_3/A_1)$ that event A_3 will occur. This will be greater than 0 and less than 1. As we get closer in the causal chain to event A_3 , then we are able to predict with a higher probability that event A_3 will occur: thus since A_2 is closer to A_3 than A_1 is, $P(A_3/A_2) > P(A_3/A_1)$. The probability of the prediction will increase towards 1 as we get closer to A_3 . So Reichenbach's basic intuition, as expressed in the first two equations, is that the closer one gets to an event in the causal chain, the higher its probability becomes.

A simple example may help to illustrate this intuition. Suppose that we are launching a missile and hope to hit a target some distance away. Knowing that the missile was launched (event A_1) certainly raises the probability of it hitting the target (event A_3). However, knowing that the missile is still on target when it crosses a certain tracking station between the launch site and the target (event A_2), enables us to infer with even higher probability that the missile will hit the target. If we were to determine that the missile was on target at a later stage in its trajectory, we would be even more confident that it would hit the target. It appears that our confidence in it hitting the target is raised because we believe that there is less that can go wrong and cause the missile to veer off course as it gets

closer and closer to the target. Reichenbach's intuition about causal betweenness seems to be supported by our experience of examples such as these.

3.1.2.2. Screening Off. The other part of Reichenbach's basic intuition is captured by equation (3.3), and is a version of the Markov property. This tells us that if we have a causal chain, an earlier event cannot affect later events except through the intermediate links. One can view this as a partial limitation on action at a distance: if there is a causal chain connecting two events, the only way the earlier event can affect the later event is through the intermediate links.

Another way of looking at this is that events that happen before A_2 are irrelevant to making a prediction of whether A_3 will occur, once we know that A_2 has occurred. Imagine that it is possible to know the complete state of a system at time t ; then states of the system before t are irrelevant to any prediction about the future of the system. Similarly, if we knew the complete state of the universe today, we could make certain predictions about the state of the universe tomorrow; Reichenbach's claim is that the state of the universe yesterday would not help us make the prediction of what the universe would be like tomorrow, once we know what the universe is like today. This idea, captured by equation (3.3), is known as the screening off relation. If the probability of A_3 given A_1 and A_2 , $P(A_3/A_1 \& A_2)$, just equals the probability of A_3 given A_2 , $P(A_3/A_2)$, then knowledge of A_1 is irrelevant to a prediction of whether A_3 will occur, once we know that A_2 has occurred. In this situation we say that A_2 screens off A_1 from A_3 ; A_2 contains all of the information that A_1 contains, and more, relevant to a prediction of the occurrence of A_3 . When an event is screened off from another event, it is no longer predictively informative about that event. This is the other basic intuition that Reichenbach was trying to capture in his definition of causal betweenness.

3.1.2.3. Causal Betweenness and the Causal Net. The relation of causal betweenness enables us to determine the structure of the causal net. Reichenbach says that if we determine that the following hold,

$$\begin{aligned} &\text{btw}(A_1, A_2, A_3), \\ &\text{btw}(A_1, A_2, A_4), \\ &\text{btw}(A_4, A_2, A_3), \end{aligned}$$

then we can know that the four events have the causal structure diagrammed in

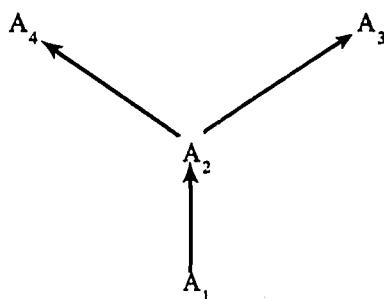


Figure 3-1: Causal Betweenness in a Causal Fork

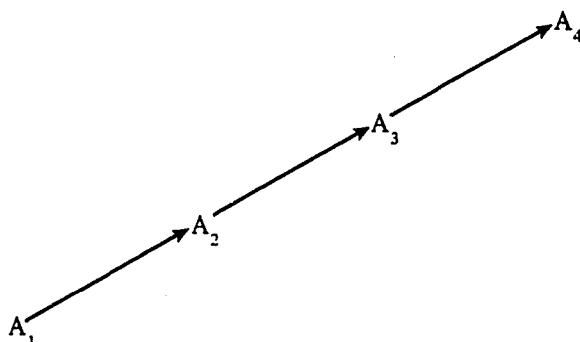


Figure 3-2: Causal Betweenness in a Causal Chain

figure 3-1 and not that of figure 3-2. The betweenness relations of forks and chains differ considerably. In the chain diagrammed in figure 3-2, $\text{btw}(A_4, A_2, A_3)$ is false, whereas it is true of the fork diagrammed in figure 3-1. Using probability relations to pick out causal relations finds its roots in this idea of Reichenbach's.

3.1.3. Critical Discussion of Causal Betweenness

Since the definition of causal betweenness formalizes the idea that the closer one gets to another event in a causal chain, the higher its probability will be, a rejection of the definition of causal betweenness will most likely involve a rejection of that principle. The only rejection of the definition of causal betweenness that I know of is due to Salmon. Salmon points out that not all events in a causal chain need be positively relevant to later events in that chain; sometimes things happen "the hard way." An example of this is given in a game of craps, in which there is more than one way to win. If a 7 or 11 is rolled on the first toss, the shooter wins automatically. If the shooter does not roll a 7, 1, 2, 3, or 12, he may win

by an alternative method. But if he rolls a 4, his probability of winning is reduced, even though his rolling a 4 may be in the causal chain leading to his winning. This shows that some events in a causal chain may be negatively relevant to a later event. Other examples like this abound. Since Reichenbach's definition of causal betweenness requires that all events in the causal chain be positively relevant to later events, we see that Reichenbach's definition is defective. We will discuss the legitimacy of this counterexample in a later chapter.

3.2. Causal Forks

3.2.1. Conjunctive Forks

Reichenbach had another discussion of probabilistic causality in connection with common causes. Reichenbach noticed that when we have a statistical correlation between some events, we should look for a common cause. In other words, if certain events occur together more often than we would expect if they were causally independent of one another, then we should look for a common cause, unless of course, one is a cause of the other. Reichenbach also claimed that knowledge of one of the effects of a common cause was irrelevant to knowledge of the other effect, once we had knowledge of the common cause. Reichenbach (1956, p. 159) formalized these ideas and assumed that all cases of a common cause C with effects A and B satisfy the following relations:

$$P(A \& B/C) = P(A/C) * P(B/C), \quad (3.4)$$

$$P(A \& B/\neg C) = P(A/\neg C) * P(B/\neg C), \quad (3.5)$$

$$P(A/C) > P(A/\neg C), \quad (3.6)$$

$$P(B/C) > P(B/\neg C). \quad (3.7)$$

Reichenbach showed that from these four assumptions we can derive the following inequality:

$$P(A \& B) > P(A) * P(B). \quad (3.8)$$

This equation tells us that the conjunction of A and B occurs more frequently than we would expect them to if they were independent of one another. For this reason, Reichenbach calls a causal fork which satisfies equations (3.4)–(3.7) a *conjunctive fork*. A conjunctive fork is a causal fork that makes the conjunction

of two or more events more probable than they would be if they were independent events. An example of a conjunctive fork, given by Salmon (1980, p. 59), concerns a common cause of illness. Suppose, in the above equations, we let

A=the illness of Smith on a certain night,
 B=the illness of Jones on a certain night, and
 C=the presence of spoiled food in their dormitory dinner that evening.

In this situation we explain the improbable coincidence that both Smith and Jones are ill by appealing to a common cause. We should notice that equations (3.4) and (3.5) tell us that the common cause C screens off A from B and that the absence of the common cause, $\neg C$, also screens off A from B. This will be important later in our discussion of probabilistic causality.

3.2.2. Conjunctive Forks and Common Causes

Salmon has presented counterexamples to show that not all conjunctive forks are common causes. Salmon (1980, p. 59) gives the following example, first presented by Ellis Crasnow:

Brown usually arrives at his office about 9:00 a.m., fixes himself a cup of coffee, and settles down to read the morning paper for half an hour before beginning any serious business. Upon occasion, however, he arrives at 8:00, and his secretary has already brewed a fresh pot of coffee, which she serves him immediately. On precisely the same occasions, some other person meets him at his office and they begin work quite promptly. This coincidence--the coffee being ready and the other person being at his office--demands explanation in terms of a common cause. As it happens, Brown usually takes the 8:30 bus to work in the morning, but on those mornings when the coffee is prepared for his arrival and the other person shows up, he takes the 7:30 bus. It can plausibly be argued that the three events, A (the coffee being ready), B (the other person showing up), and C (Brown taking the 7:30 bus), satisfy Reichenbach's requirements for a conjunctive fork. Clearly, however, Brown's bus ride is not a cause either of the coffee being made or the other person's arrival. The coincidence does, indeed, require a common cause, but that event is a telephone appointment made by the secretary on the previous day.

Salmon notes that counterexamples like this can be constructed anytime we have an effect of the common cause that is strictly correlated with the common cause; another way of saying that is to notice that counterexamples will arise any time the

common cause is a necessary and sufficient cause of an event that is temporally prior to the two events for which we seek an explanation. These counterexamples show that events A, B, and C can satisfy the definition of a conjunctive fork, and yet C may not be a common cause of A and B. The counterexamples discussed above arise because there may be a D, which is necessary and sufficient for C, as well as the two events A and B.

At this point we should realize that the above counterexamples are counterexamples to the claim that all events A, B, C that satisfy equations (3.4) through (3.7) are cases in which C is a common cause of A and B. However it is clear that Reichenbach himself did not believe that all A, B, and C which satisfied equations (3.4) through (3.7) were cases in which C was a common cause of A and B; Reichenbach noticed that it was possible for C to be a common effect instead of a common cause. Reichenbach simply assumed that in cases in which C was a common cause of A and B, equations (3.4) through (3.7) would be satisfied by A, B, and C; he claimed that C being a common cause of A and B was a sufficient condition, but not a necessary condition, of equations (3.4) through (3.7) being satisfied.

3.2.3.-Interactive Forks

Although we have seen that satisfying the definition of a conjunctive fork is not a sufficient condition of C being a common cause of A and B, it may be true that satisfying the definition of a conjunctive fork is a necessary condition of C being a common cause of A and B. In other words, perhaps all cases involving a common cause satisfy equations (3.4) through (3.7). A problem arises for this because of the existence of what are called interactive forks. Salmon (1978) originally defined an interactive fork as one which satisfied the following conditions:

$$P(A \& B/C) > P(A/C) * P(B/C), \quad (3.9)$$

$$P(A \& B/\neg C) = P(A/\neg C) * P(B/\neg C), \quad (3.10)$$

$$P(A/C) > P(A/\neg C), \quad (3.11)$$

$$P(B/C) > P(B/\neg C). \quad (3.12)$$

The definition of an interactive fork differs from that of a conjunctive fork in that the equality in equation (3.4) is replaced by the inequality in (3.9). In an interactive fork one prong of the fork may be a better indication of the state of

the other prong of the fork than the common cause is; the common cause does not screen off one prong from the other prong.

Salmon (1978, pp. 692–694) gives the following example of an interactive fork:

If, for example, an energetic photon collides with an electron in a Compton scattering experiment, there is a certain probability that a photon with a given smaller energy will emerge, and there is a certain probability that the electron will be kicked out with a given kinetic energy.... However, because of the law of conservation of energy, there is a strong correspondence between the two energies--their sum must be close to the energy of the incident photon. Thus, the probability of getting a photon with energy E_1 and an electron with energy E_2 , where $E_1 + E_2$ is approximately equal to E (the energy of the incident photon), is much greater than the product of the probabilities of each energy occurring separately. Assume, for example, that there is a probability of 0.1 that a photon of energy E_1 will emerge if a photon of energy E impinges on a given target, and assume that there is a probability of 0.1 that an electron with kinetic energy E_2 will emerge under the same circumstances (where E , E_1 , and E_2 are related as the law of conservation of energy demands). In this case the probability of the joint result is not 0.01, the product of the separate probabilities, but 0.1, for each result will occur if and only if the other does.

In this example the common cause does not screen off one prong of the fork from the other prong. The existence of interactive forks shows that not all cases in which C is a common cause of A and B are conjunctive forks: some of them are interactive forks. Interactive forks show that being a conjunctive fork is not a necessary condition of being a case involving a common cause. The proper way to view Reichenbach's definition of a conjunctive fork is not as a necessary or sufficient condition for common causes, but rather as a certain type of case involving a common cause. It is a fact about our world that many cases of common causes are conjunctive forks and satisfy relations (3.4) through (3.7).

3.3. Causal Relevance

3.3.1. Causal Relevance and Mark Transmission

Reichenbach was perhaps the first philosopher to notice the close connection between causal sequences and the ability to transmit a mark. Basically, causal sequences can transmit a mark, whereas pseudo-processes cannot transmit a mark. A mark is simply a modification of the sequence that appears in later parts of the sequence. We can then easily see that if a process is not a genuine causal process, a mark will not be transmitted and show up at later stages of the process. For example, consider a shadow of a car that moves along as the car moves. Clearly the shadow is a pseudo-process and not a genuine causal process. Any mark that is placed on the shadow of the car will not appear in any later stages of the shadow. This is different from placing a mark on the car, because a mark on the car will appear at later times on the car. Only genuine causal processes are able to transmit marks.

Reichenbach (1956, p. 290) defines a relation of *causal relevance*:

Definition 2: If a mark made in an event A_i shows in an event A_k , then A_i is causally relevant to A_k .

When we say that A_i is causally relevant to A_k , we are saying that A_i "is a cause contributing to the existence of A_k ." This definition reflects Reichenbach's belief that the marking process can pick out causal relations.

3.3.2. Causal Relevance and Probability Relations

3.3.2.1. Assumption Alpha. Reichenbach believes that there are certain relations between the causal relations established by the marking process and the probability relations among events. He investigates this correspondence by making various assumptions. The first assumption (1956, p. 201) is:

Assumption α :

If a mark made in A_i shows in A_k , then

$$P(A_k/A_i) > P(A_k).$$

This assumption is a necessary condition; it is not a sufficient condition also. That

this probability relation holds whenever a mark can be transferred is an empirical assumption that is justified by our experience. In our experience we find that events that are causally relevant are also positively relevant in the statistical sense.

3.3.2.2. Assumption Beta. Reichenbach's second assumption (1956, p. 201) is:

Assumption β :

If a mark is made in A_i , then either

$$P(A'_k/A'_i) = P(A_k/A_i) \quad (3.13)$$

or

$$P(A'_k/A'_i) = P(A_k/A_i) \quad (3.14)$$

We denote by A' the event A that is also marked. Assumption β essentially says that the marking process does not change the existing probabilities. If (3.13) holds, we say the the mark was transferred from A_i to A_k ; if (3.14) holds, we say that the mark was not transferred from A_i to A_k . Reichenbach notes that if A_k temporally precedes A_i , then (3.14) will hold, because a mark cannot travel backwards in time. (3.14) can also hold if the mark disappears before it reaches A_i . Some marks are less permanent than others, and even if two events are causally relevant there may not be a mark transferred between them. Reichenbach is claiming that whether or not the mark reaches the later event the causal relations are unchanged by the marking process.

3.3.2.3. Assumption Gamma. Reichenbach's third assumption (1956, p. 202) deals with the problem of marks that disappear:

Assumption γ :

If A_2 screens off A_1 from A_3 , and if a mark made in A_1 shows in A_3 , then it also shows in A_2 .

This assumption tells us that if a causal process is marked and the mark disappears somewhere along the line, then the mark will not reappear somewhere later in the causal process. In other words, a causal process is continuous in the sense that marks do not suddenly disappear and appear. Reichenbach considers this principle to be based on the principle of action by contact. It too is a principle that is justified by our experience.

3.3.2.4. Assumption Delta. Reichenbach (1956, p. 204) then presents a more general version of assumption γ :

Assumption δ :

If a set $A_2^1 \dots A_2^n$ screens off A_1 from A_3 , and if a mark made in A_1 shows in A_3 , then it also shows in at least one of the events $A_2^1 \dots A_2^n$.

Reichenbach generalized assumption γ because sometimes an earlier event will not screen off a certain event, but a conjunction of earlier events will permit screening off.

3.3.2.5. Probabilistic Definitions of Causal Relevance. Given assumptions α , β , and δ , Reichenbach (1956, p. 204) develops another definition of causal relevance:

Definition 3: An event A_1 is causally relevant to a later event A_3 if $P(A_3/A_1) > P(A_3)$ and there exists no set of events $A_2^1 \dots A_2^n$ which are earlier than or simultaneous with A_1 such that this set screens off A_1 from A_3 .

This definition is based on assumption δ . We could also give a less general definition of causal relevance that is based on assumption γ :

Definition 4: An event A_1 is causally relevant to a later event A_3 if $P(A_3/A_1) > P(A_3)$ and there exists no A_2 which is earlier than or simultaneous with A_1 such that A_2 screens off A_1 from A_3 .

Since definition 4 is easier to work with than definition 3, we will base our discussion of causal relevance upon it. Anything we assert about definition 4 could be generalized and applied to definition 3. We should notice that definition 4 does not make reference to a marking process, and may be useful where it is difficult to mark a process. We should also notice that this definition presupposes a temporal direction; thus it cannot be used to define a direction of time.

3.3.3. Causal Betweenness, Conjunctive Forks, and Causal Relevance

A comparison of the definition of causal betweenness with the probabilistic definition of causal relevance will quickly reveal that they are different definitions, and not merely the same definition stated in two different ways. Suppose that A_1 is causally relevant to A_2 and that A_2 is causally relevant to A_3 . It is clear that in

this case A_2 is causally between A_1 and A_3 . But merely from the fact that A_1 is causally relevant to A_2 and that A_2 is causally relevant to A_3 , we cannot derive the three equations which define causal betweenness. The definition of causal relevance in no way implies the Markov property that is a part of the definition of causal betweenness. Even if we add the Markov property to the situation in which A_1 is causally relevant to A_2 and A_2 is causally relevant to A_3 , we cannot derive equations (3.1) and (3.2). For example, in equation (3.1) we could easily get $1 > P(A_3/A_2)$ and $P(A_3) > 0$. If we assume that the chain $A_1 \rightarrow A_2 \rightarrow A_3$ is a two state chain with the Markov property, we can derive $P(A_3/A_1) > P(A_3)$, so we can prove that $1 > P(A_3/A_2)$ and $P(A_3/A_1) > P(A_3) > 0$. However, it does not appear possible to prove that $P(A_3/A_2) > P(A_3/A_1)$. Thus we can see that an event A_2 may be causally between events A_1 and A_3 according to the definition of causal relevance, and yet we cannot derive that it is also causally between A_1 and A_3 according to our probabilistic definition of causal betweenness. This shows that in cases in which these two definitions should give equivalent results, if they do it is because of an accidental fact about the world and not because of any logical necessity.

Similar results can be proven concerning conjunctive forks. When we have a conjunctive fork, neither of the prongs A and B can be causally relevant to the other according to definition 3, because there is an earlier event that screens it off from the other prong. This is a desirable result. But, given the definition of a conjunctive fork, I am unable to prove that C is causally between A and B. The problem is in proving that $P(A/C) > P(A/B)$. It appears as if this is another case in which the definitions overlap, and the overlap is not due to the probability calculus.

3.3.4. The Equivalence of Definitions 2 and 4

Reichenbach believes that definitions 2 and 3, which define causal relevance, are extensionally equivalent. We have observed that whenever an event is causally relevant according to definition 4, that it is possible to construct a mark proceeding from the cause to the effect. This is not logically necessary: it is just an empirical fact about the world. The inference the other way is based on the three assumptions given. Reichenbach reasons as follows. Suppose that A_1 is causally relevant to A_3 according to definition 2; this means that it is possible to transmit a mark from A_1 to A_3 . From assumption α we can derive that $P(A_3/A_1) > P(A_3)$.

This is half of the definition that we want. Now suppose that A_2 is some event that is earlier than or simultaneous with A_1 . We can then derive that A_2 cannot screen A_1 off from A_3 . This is done by noticing that assumption γ claims that if A_2 screened A_1 off from A_3 , then a mark made in A_1 that showed in A_3 would also have to appear in A_2 . But the mark cannot appear in A_2 , because a mark always travels forward in time; a mark made in A_1 cannot appear in any event that is earlier than or simultaneous with A_1 . Thus assumption γ tells us that A_2 cannot screen off A_1 from A_3 . But this is just Reichenbach's requirement that there exist no A_2 that is earlier than or simultaneous with A_1 such that A_2 screens off A_1 from A_3 . Thus given Reichenbach's assumptions, definition 2 entails definition 4.

On first examination it appears as if Reichenbach did not make use of assumption β , and that it is not needed to demonstrate the equivalence of definitions 2 and 3. However a closer look reveals that assumption β is used, though indirectly. Reichenbach attempted to intuitively motivate his three assumptions by giving reasons to accept them. He thought that we should accept assumption γ , because assumption γ is a necessary condition of the assumption that the marking process does not disturb the various screening off relations. In other words, if assumption γ is not true, it can be shown that the marking process destroys the screening off relations that might have held.

Reichenbach's proof is as follows. Let us assume the negation of assumption γ and deduce a contradiction. Thus we assume that a mark made in A_1 appears in A_3 but not in A_2 . He also assumes that the marking relation does not disturb the screening off relations. This principle he expresses by the following probability relations:

$$P(A'_3/A'_1 \& A'_2) = P(A'_3/A'_1) \quad (3.15)$$

$$P(A'_3/A'_1 \& A'_2) = P(A'_3/A'_2). \quad (3.16)$$

These probability relations express the idea that the screening off relation holds for events that are marked and for events that are unmarked. Reichenbach's first step in the proof is to use the theorem of total probability to get:

$$P(A'_3/A'_1) = P(A'_2/A'_1)P(A'_3/A'_1 \& A'_2) + P(-A'_2/A'_1)P(A'_3/A'_1 \& -A'_2).$$

Reichenbach then uses assumption β to replace $P(A'_3/A'_1)$ by $P(A'_3/A_1)$, and to replace $P(A'_2/A'_1)$ by $P(A_2/A_1)$ in the above equation (see lines (3.13) and (3.14),

respectively). When he combines the previous replacements with equation (3.16) he gets:

$$P(A_3/A_1) = P(A_2/A_1)P(A'_3/A_2) + [1 - P(A_2/A_1)]P(A'_3/A_1 \& \neg A_2).$$

We must question whether this is a legitimate use of assumption β . Assumption β tells us that either $P(A'_3/A'_1) = P(A_3/A_1)$ or that $P(A_3/A'_1) = P(A_3/A_1)$; we know that one of the disjuncts is true, but we don't know which one. It appears as if Reichenbach simply assumed that the first disjunct is true, which is an unjustified assumption. If we look at assumption β , we see that the second disjunct was for cases in which A_3 was earlier than or simultaneous with A_1 , or for cases in which the mark disappeared. Thus if we assume that A_1 is earlier than A_3 and that the mark does not disappear, then it looks as if we are justified in assuming that the first disjunct is true in this case. This is a strong assumption, but it is necessary for the proof.

Similarly, assumption β tells us that either $P(A'_2/A'_1) = P(A_2/A_1)$ or that $P(A_2/A'_1) = P(A_2/A_1)$; again we do not know which of the disjuncts is true in this case. It appears as if Reichenbach assumed that the second disjunct was true, but this is an unjustified use of assumption β . Recalling Reichenbach's remarks about assumption β , we would be justified in using the second disjunct if either A_2 is earlier than A_1 or if the mark disappears before it reaches A_2 . In order for this proof to be successful, we must then assume that either A_2 is earlier than A_1 or that the mark from A_1 disappears before it reaches A_2 .

Summarizing our additional assumptions, we find that we must assume that A_1 is earlier than A_3 , that the mark does not disappear from A_1 to A_3 , and that either A_2 is before A_1 or that the mark disappears from A_1 to A_2 . It will profit us to inquire as to what section of the causal net is compatible with these further assumptions. One possibility might be a causal fork as in figure 3-3.

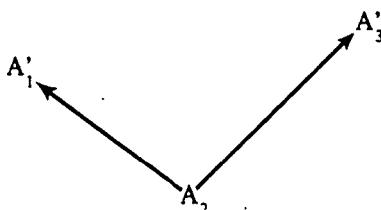


Figure 3-3: Marks in a Causal Fork

In figure 3-3, we have A_1 earlier than A_3 , and A_2 earlier than A_1 , which is what the assumptions require. The problem with a causal fork is that the assumptions also require that the mark not disappear from A_1 to A_3 . If this is a causal fork, and there is no causal connection between A_1 and A_3 , it does not seem possible that a mark could be transmitted from A_1 to A_3 . Thus the causal fork of figure 3-3 does not seem to satisfy all of our additional assumptions.

Another possibility might be a causal chain as diagramed in figure 3-4. In figure 3-4, we have A_1 before A_3 , A_1 before A_2 , the mark does not disappear from A_1 to A_3 , and the mark does disappear from A_1 to A_2 .

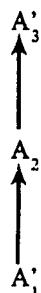


Figure 3-4: Marks in a Causal Chain

If this ever did occur it would be an amazing situation. We commonly think that if a mark disappears from a causal process, it will not reappear at a later stage of the process, and indeed this is the very situation that Reichenbach is trying to exclude with assumption γ . Although this situation is compatible with our additional assumptions, Reichenbach believes that it is incompatible with the assumption that A_2 screens off A_1 from A_3 and that the marking process does not disturb the screening off relation. Furthermore, we find that this situation is incompatible with assumption γ , which is what we are trying to motivate. The causal chain is compatible with all of the assumptions of assumption γ , but it is not compatible with them and the consequent of assumption γ . The consequent of assumption γ claims that the mark also shows in A_2 , which contradicts the assumption of figure 3-4 that the mark does not show in A_2 . Consequently we must exclude this situation from the possible sections of the causal net that are compatible with our assumptions.

It appears to me as if the only real possibility left is that of the causal chain diagramed in figure 3-5. In figure 3-5 we have A_1 earlier than A_3 , the mark does



Figure 3-5: Causal Chain In Which Mark Appears

not disappear from A_1 to A_3 , and A_2 is earlier than A_1 . This is certainly a possible situation, but it is strange that A_2 screens off A_1 from A_3 , as is required in the hypothesis of assumption γ . This situation would occur if $P(A_1/A_2)=1$ and $P(A_1 \& \neg A_2)=0$. But Reichenbach does not want to allow the limiting values of one and zero to be legitimate, as we saw in his discussion of causal betweenness. We also noticed that Reichenbach believed that the closer one came to an event in the causal chain, the higher the probability of the event became. The limiting value of one is approached as one moves closer to the event, although it never reaches one. The important idea is that by looking at an event closer to another in a causal chain, our probabilities will always rise. Reichenbach (1956, p. 189; 1949 p. 237) says that this is a relationship which holds in general for continuous probability sequences. However, the causal chain diagrammed in figure 3-5 does not fit this requirement, and I can think of no example of which it does fit. This appears to be an example which is consistent with the probability calculus, but which does not seem to have any causal analogues. Further examples like this will arise when we discuss Suppes' criticisms of a definition similar to definition 4. Thus I conclude that it is unclear as to whether there are any physical causal situations in which all of the assumptions needed so far are true. It appears doubtful to me that there are any.

Reichenbach's (1956, p. 202) next step in his proof that assumption γ is necessary if we want the marking process to leave the screening off relations undisturbed is to claim that since A_3' can occur only if A_3 does not occur, the following holds:

$$P(A_3'/A_2) \leq 1 - P(A_3/A_2). \quad (3.17)$$

We must question whether it is true whether events A_3 and A_3' are disjoint, or more

generally, whether a marked event can occur only if the unmarked event does not occur. It seems to me that there is no obviously correct answer to this question, and we can define events however we want to, as long as we do it consistently. Thus we can claim that events and marked events are disjoint. However, if we follow Reichenbach in doing this, it seems as if assumption β is no longer obviously true.

The first disjunct of assumption β , that $P(A'_k/A'_i) = P(A_k/A_i)$, seems true in the situations that Reichenbach thought it was true; if A_k is later than A_i and the mark does not disappear, this probability relation seems correct. However, the second disjunct, $P(A_k/A'_i) = P(A_k/A_i)$, does not seem to be true in the cases that Reichenbach thought it was true. Let us suppose that A_k is earlier than A_i : what are our reasons for believing that $P(A_k/A'_i) = P(A_k/A_i)$? I see no reason to believe that the probability that a cause is unmarked given that the effect is marked is equal to the probability that the cause is unmarked given that the effect is unmarked. It would appear that the probability that a cause is unmarked given that the effect is marked might be much lower than the probability that the cause is unmarked given that the effect is unmarked; this would be the case if we knew that the mark did not disappear often. We should notice that although this disjunct of assumption β is unreasonable upon the assumption that a marked event and an unmarked event are disjoint, it does not seem unreasonable upon the assumption that they are not disjoint. Let us suppose that A'_i is a proper subset of A_i . We could define A'_i by saying that an event x is a member of A'_i if and only if x is a member of A_i and x is a member of M , where M is the class of marked events. This would give us the result that the class of marked events is a subset of the class of events in question. Under this interpretation of events, it would seem reasonable that $P(A_k/A'_i) = P(A_k/A_i)$. This says that the probability that a cause occurred given that a marked effect occurred is equal to the probability that a cause occurred given that the effect occurred, which seems correct. This probability relation would hold as long as the marking process did not disturb the existing probability relations. Thus we see that in order for the second disjunct to be true when A_k is earlier than A_i , we must interpret events in a manner differently than Reichenbach interpreted them in assumption γ .

Now let us suppose that sometimes the mark disappears during the

transmission from A_i to A_k ; in this situation the second disjunct, $P(A_k/A'_i)=P(A_k/A_i)$, is supposed to be true. But if an event and the same event marked are considered disjoint, I see no reason to believe that that is true; on the contrary, it appears false. For example, suppose the mark disappears about half of the time; it would then seem like $P(A_k/A'_i)=1/2P(A_k/A_i)$. In general, if the mark disappears a certain percentage of the time, it will almost always be false that $P(A_k/A'_i)=P(A_k/A_i)$. Reichenbach gives an example in which snow on the roof of a train melts as time goes on; he was well aware that a mark can disappear as time goes on. But by requiring that $P(A_k/A'_i)=P(A_k/A_i)$ we force the marking event to change the existing probability relations, which is just what assumption β was meant to rule out.

We could remedy this difficulty by replacing the second disjunct of assumption β with:

$$P(A_k \vee A'_k/A'_i) = P(A_k/A_i). \quad (3.18)$$

This equation tells us that if a cause is marked, then the probability of getting either a marked effect or an unmarked effect is the same as the probability of getting an unmarked effect given an unmarked cause. This is essentially saying that the marking relation does not change the existing probabilities. Equation (3.18) also seems to hold when A_k is earlier than A_i , which was the other purpose of equation (3.14). Given a marked effect, the probability of having either a marked cause or an unmarked cause is equal to the probability of having an unmarked cause given an unmarked effect. This too seems to require that the marking process leave the existing probability relations undisturbed. So if we wish to interpret events in such a way that an event and that event marked are disjoint events, we will have to replace equation (3.14) in assumption β by equation (3.18).

Even though it is possible to replace equation (3.14) by equation (3.18), it seems preferable to me to interpret events in such a way that both an event and that event marked can occur: they are not disjoint. We have already seen that this interpretation solves the problems facing equation (3.14) when A_k is earlier than A_i . It also seems to be true when A_k is later than A_i and the mark disappears part of the time. If A_k and A'_k are not disjoint, and A'_k is a proper subset of A_k , then it certainly seems true that $P(A_k/A'_i)=P(A_k/A_i)$, when the mark disappears part of the time. This will be true as long as the mark does not disturb the existing probability relations, which is what we want.

We thus find ourselves faced with a dilemma. If we decide that both an event and the same event marked are disjoint events and cannot both occur, then our proof fails because assumption β then appears to be false. However if we interpret events in such a way that both an event and the same event marked can occur, then the proof fails because it required that they both not occur. It appears that the proof fails no matter which way we interpret events. It also failed because of the improper use of assumption β that we discussed earlier. Thus we must conclude that Reichenbach's attempt to show that assumption γ is necessary in order to claim that the marking process does not disturb the screening off relation is unsatisfactory.

Even though the above proof failed, assumption γ still has some intuitive appeal, and seems to be correct in many cases. The only instance I can think of in which it is false deals with interactive forks. Suppose we have an interactive fork as diagrammed in figure 3-6.

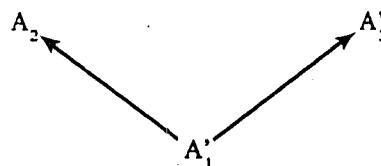


Figure 3-6: Marking in an Interactive Fork

We will suppose that A_2 screens off A_1 from A_3 . However, it seems possible to me that a mark made at A_1 may be transmitted to A_3 without being transmitted to A_2 . Consider the following example of Salmon's. Suppose that A_1 is one billiard ball colliding with another, A_2 is one of the billiard balls falling in the corner pocket, and A_3 is the other billiard ball falling in the other corner pocket. In this case it is plausible to suppose that A_2 screens off A_1 from A_3 . Now suppose that the billiard ball that collided with the other is marked at or before A_1 . It seems reasonable that the ball will still have the mark at A_3 , and yet it does not seem necessary that the mark have been passed on to A_2 . This situation is possible because two causal processes can interact, and yet not all marks on one will be passed on to the other. We certainly do not want to claim that a mark made at A_1 cannot be transmitted to both A_2 and A_3 ; we only want to claim that it does not have to be transmitted to both. Thus assumption γ appears to be false in this situation.

3.3.5. Modifications of Assumption Gamma

However, the falsity of assumption γ does not imply that Reichenbach was wrong in thinking that definition 2 with the proper assumptions implies definition 4. Perhaps we can formulate a weaker version of assumption γ that will enable us to derive definition 4 from definition 2, and yet not fall prey to any counterexamples. Let us consider the following proposal:

Assumption γ^* :

If A_2 screens A_1 off from A_3 , and if A_2 is earlier than A_3 and later than A_1 , then if a mark made at A_1 shows at A_3 , it must show at A_2 .

This appears to be what Reichenbach meant by assumption γ , and it fits the diagram of figure 3-4. If we have a causal chain, it seems perfectly reasonable that if the mark is stopped at some place, it will not reappear later. But the problem is that assumption γ^* also fits the interactive fork diagrammed in figure 3-6, as long as we make sure that A_2 is earlier than A_3 . So assumption γ^* is also false, for the same reason that assumption γ is. However, even if assumption γ^* were true, it could not be used to prove that definition 2 implies definition 4. In proving definition 4, we used assumption γ to claim that if there were any event earlier than A_1 which screened off A_1 from A_3 , then the mark made in A_1 would have to show in A_2 if it showed in A_3 . But in this case A_2 is before A_1 , so we could not have been able to use assumption γ^* even if it had been true.

Let us take the required temporal relations into account with the following proposal:

Assumption γ^{} :**

If A_2 screens A_1 off from A_3 , and if A_2 is earlier than both A_1 and A_3 , then a mark made at A_1 that shows at A_3 must also show at A_2 .

The temporal relations in assumption γ^{**} are compatible with either the causal fork diagrammed in figure 3-3, or the causal chain diagrammed in figure 3-5. I have already mentioned that if we exclude the limiting values of zero and one from our probability relations, then figure 3-5 becomes a very puzzling case. If the causal chain in figure 3-5 is a genuine physical possibility, definition 4 is faulty. In figure 3-5, we would certainly claim that A_1 is causally relevant to A_3 , and yet it would not be causally relevant according to definition 3, because there is an earlier event A_2 which screens it off from A_3 . Assumption γ^{**} appears false when

applied to cases like figure 3-5, but that need not bother us, because definition 3 is also inadequate when applied to cases like that in figure 3-5. Until an example is presented which satisfies the relations described in figure 3-5, I think it best to ignore this case. It seems to be a fact about the world that cases of this sort do not occur.

This leaves figure 3-3 as the situation described in assumption γ^{**} . In this case it certainly seems true that if a mark is transmitted from A_1 to A_3 , it must also show in A_2 . Since there is no causal connection between A_1 and A_3 , no mark could be transmitted between them unless it went through A_2 . Another way to look at it is to realize that no mark could possibly be transmitted from A_1 to A_3 , since there is no causal connection between them. But if no mark can go from A_1 to A_3 , then it certainly seems true that if any mark does go from A_1 to A_3 it would also appear in A_2 . The only reason for accepting assumption γ^{**} is the belief that a mark cannot go from A_1 to A_3 in the situation described. The reason we believe that a mark cannot go from A_1 to A_3 is because we believe that we have a causal fork, and not a chain, when the probability relations given are true. This leads us to notice the intimate connection between assumption γ^{**} and definition 4. It appears that they stand or fall together. If one rejects assumption γ^{**} , then he would also reject definition 4. And if one rejects definition 4, he will also reject assumption γ^{**} , unless he rejects assumption α , which is unlikely. We see that assumption γ^{**} is a very strong assumption; it would be nice to have a proof that if it did not hold, that a marking process would disturb the screening off relation, but I do not see how to give such a proof.

To summarize our discussion so far, we can conclude that we can derive definition 4 from definition 2, assumption α , and assumption γ^{**} . We no longer need assumption β , because it was just used in the proof to make assumption γ seem more plausible. Since assumptions α and γ^{**} are known by experience, we must claim that the equivalence of definitions 2 and 4 is one that we have observed in our experience of the world, and is not known *a priori*.

3.3.6. Problems With Definition 3

Reichenbach's definition of causal relevance is a major contribution to the discussion of probabilistic causality, and all later theories will have some modified version of his definition of causal relevance. Thus we must investigate the shortcomings of this definition.

3.3.6.1. The Occurrence of The Cause and Effect. Some have objected to definitions of causal relevance that are similar to definition 3 on the following grounds. The claim is that such definitions are too broad, or do not present sufficient grounds for calling an event a cause because there will be times in which the cause occurs but the effect does not occur. Suppose that someone throws a baseball at a window, but just before it hits the window someone jumps up and catches it. It appears that throwing the baseball at the window is a probabilistic cause of the window breaking: it raises the probability of the window breaking and there doesn't seem to be any earlier event that screens it off. But in this case we might want to say that the cause is defeated; something intervened to prevent the cause from bringing about its effect. If we were giving an analysis of deterministic causation, this example would certainly be a problem, because the cause was defeated. But it is not clear that this is a counterexample to a theory of probabilistic causality. The purpose of a theory of probabilistic causality is to account for cases in which the cause does not always produce its effect. Thus in this example we might not have an example of a defeater, because a probabilistic cause is not defeated just because the cause occurs and the effect does not occur. It is not clear what a defeater for a probabilistic cause would be; it would certainly be different than in the deterministic cases.

One might object to saying that the throwing of the baseball at the window was a probabilistic cause of the window breaking on the grounds that in order to say that A is a probabilistic cause of B, both A and B must occur. Suppes (1970, p. 37) discusses an example given by Deborah Rosen:

Accepting the current data about the relationship between smoking and lung cancer, we might very well be led to say that John's smoking three packs a day of unfiltered cigarettes in his teen years is a *prima facie* cause of his getting cancer at the age of 60, but then John may die prematurely in his twenties as the result of an automobile accident. Our causal statement then seems peculiar in terms of ordinary distinctions....

Both Suppes and Rosen find it puzzling that A could be a probabilistic cause of B even if B failed to occur. This sense of puzzlement might even be increased if we look at cases in which the cause does not occur. Suppose that John has never smoked a cigarette in his life, but unfortunately develops lung cancer at the age of 60. It would certainly seem strange to say that a cause of his developing cancer was smoking three packs a day of unfiltered cigarettes in his teen years. A similar situation would occur if John neither smoked three packs a day of unfiltered cigarettes in his teen years, nor ever contracted lung cancer. In this situation, where neither the cause or effect occur, it is unusual to claim that John's smoking three packs a day of unfiltered cigarettes as a teenager causes his lung cancer later on in life.

However, these difficulties do not arise for Reichenbach's second definition of causal relevance. Definition 3 begins by saying "an event A_1 is causally relevant to a later event A_3 if...." This implies that both A_1 and A_3 actually occur. By requiring that the cause and effect both occur, the problems mentioned in the previous paragraph are avoided.

Even though we have avoided some problems by noticing that Reichenbach required that both the cause and effect occur, one can imagine the following objection being presented. One might object that since in order to be a probabilistic cause, both the cause and effect must occur, the probabilistic cause is always followed by its effect. If this were not true, it would not be a probabilistic cause. But if a probabilistic cause is always followed by its effect, it would seem like it is also a deterministic cause. Thus it might be claimed that if we require that both a probabilistic cause and its effect occur in order to be a case of probabilistic causation, then all cases of probabilistic causation will become cases of deterministic causation. This would have the unfortunate result that there would be no probabilistic causes.

However this objection rests upon confusion. When we say that A is a probabilistic cause of B, we must remember that A and B are classes. It would be more precise to say that x_i which is a member of class A is a probabilistic cause of y_j which is a member of class B (Reichenbach 1956, p. 188). It is true that if x_i is a probabilistic cause of y_j that both x_i and y_j must occur. But that does not mean that every x_j that is a member of A is followed by a y_j that is a member of B;

this is what would be required to say that A is a deterministic cause of B. Another way of saying this is to claim that if one specific event is a probabilistic cause of another event, then they both must occur; but this does not mean that every event that is similar to the probabilistic cause will be followed by an event that is similar to the probabilistic effect. We can thus see that the above objection rests upon confusing events and classes of events.

Even though we require that both a probabilistic cause and its effect exist in order to say that we have a case of probabilistic causation, there do appear to be cases in which this does not seem appropriate. Sometimes we speak very generally and say that driving while intoxicated causes accidents. A statement like this is not claiming that a specific event in space and time is the probabilistic cause of another specific event in space and time. On the contrary, that statement is about classes of events and not the events themselves. One class of events can be called the cause of another class of events; this has no reference to specific events. When claiming that one class of events is a cause of another class of events, we do not require that any specific events exist, because the causal claim is not primarily about any specific events. In a later chapter we will return to the discussion about whether causal statements are primarily about events or types of events.

Although Reichenbach's probabilistic definition of causal relevance is not affected by the above criticisms, there are problems that definition 4 must face. One problem arises because there are events that are classified as being causally relevant according to definition 4, but none the less are not in the appropriate causal chain. The notion of causal relevance in definition 4 is too broad. Suppose we have a situation where two marksmen each fire a shot at a balloon simultaneously. Let:

- A= marksman A fires at the balloon,
- B= marksman B fires at the balloon, and
- C= the balloon breaks.

Clearly both A and B are causally relevant to C according to definition 4; each of them raises the probability that the balloon will break, and neither of them is screened off by any earlier event. Now suppose that marksman A hits the balloon and breaks it, while marksman B misses the balloon. In this case marksman B's firing at the balloon is not in the causal chain of the balloon breaking; thus it

could not be causally relevant according to definition 2, because no mark could have been transmitted from B to C. This is not to deny that a mark could have gone from B to C if marksman B had not missed the balloon. Surely we agree that B might have been a cause of C, but in this case it wasn't a cause. This example shows that an event can be causally relevant to another event according to definition 4 and not according to definition 2; the two definitions are not equivalent. This example also shows us that definition 4 is too broad: some events it classifies as causally relevant are not really causally relevant.

3.3.6.2. Suppes' Criticism of Definition 4. The question now arises as to whether definition 4 is also too narrow; are there events that are really causally relevant to other events, and yet are not classified as being causally relevant according to definition 4? In this context, we must remember that by causal relevance, Reichenbach (1956, p. 199) means "a cause contributing to the existence of" the effect.

Although Suppes never mentions Reichenbach, he does initially propose a definition of probabilistic causality that is almost identical to definition 4. The main difference is that Suppes requires that there be no event A_2 *earlier than* A_1 which screens off A_1 from A_3 , while Reichenbach required that there be no event A_2 *earlier than or simultaneous with* A_1 which screens off A_1 from A_3 . Suppes (1970, p. 22) discussed the possibility of letting A_2 be simultaneous with A_1 , but he rejected it because it resulted in the fact that there were no genuine probabilistic causes. If we let the event A_2 be the same as the event A_1 , then it is always true that there exists an event A_2 which is earlier than or simultaneous with A_1 , and which screens off A_1 from A_3 . Thus nothing would ever be causally relevant to anything else. This may seem like a trick, because an event will always screen itself off from another event. The natural response is to require that A_2 be different from A_1 , in which case the above claim that there will always be an A_2 which renders A_1 not causally relevant is unfounded.

Suppes then shows that even if we require that A_2 be different from A_1 , if we allow A_2 to be simultaneous with A_1 , the problem still arises. As long as there is an event, say A_4 , which is distinct from A_1 and such that $P(A_1 \& A_4) > 0$, there will always be an A_2 which is simultaneous to A_1 and screens off A_1 from A_3 : simply let $A_2 = A_1 \cap A_4$. In this case, A_2 screens off A_1 from A_3 because:

$$P(A_3/A_1 \& A_2) = P(A_3/A_1 \& A_1 \cap A_2) = P(A_3/A_1 \cap A_2) = P(A_3/A_2).$$

Thus Suppes concludes that A_2 cannot be simultaneous with A_1 ; this is a minor modification to Reichenbach's theory and presents no major difficulty.

Suppes (1970, p. 22) then considers a situation in which A_1 is a better predictor of A_3 than both A_1 and A_2 together are. Suppose A_2 screens off A_1 from A_3 , $P(A_3/A_1 \& A_2) = P(A_3/A_2)$, and yet A_1 is a better predictor of A_3 than A_2 is, $P(A_3/A_1) > P(A_3/A_1 \& A_2)$. It appears that these two conditions are consistent, but Suppes thinks that in this case it would be unintuitive to deny that A_1 is causally relevant to A_3 . If A_1 alone predicts A_3 with a higher probability than the joint event $A_1 \& A_2$ does, it seems as if A_1 is causally relevant to A_3 , even if it is screened off by A_2 . Suppes does not give us a concrete example which illustrates this situation, because if we knew A_1 had occurred, we could predict the occurrence of A_3 with higher probability if we didn't know that A_2 had occurred. I find this situation puzzling.

To conclude our discussion of Reichenbach, we see that Reichenbach presented probabilistic definitions of causal relevance, causal betweenness, and conjunctive forks. These definitions will be incorporated into later author's works on probabilistic causality to a very large degree. We will also find that many of the problems that Reichenbach's definitions faced will arise again to plague other theories also.

CHAPTER 4

SUPPES' PROBABILISTIC THEORY OF CAUSALITY

We have seen that Suppes believes that we need a theory of probabilistic causality in order to capture the way we ordinarily speak of causal relations. Suppes gives us a theory designed for this need in which probability relations define causal relations. The main problems facing a probabilistic theory of causality are those of distinguishing between genuine and spurious causes as well as direct and indirect causes. Suppes presents several definitions of different types of causes in an attempt to capture the distinction between genuine and spurious causes, and direct and indirect causes. It is my claim that Suppes' definitions fail to distinguish among genuine and spurious causes and direct and indirect causes. To support this claim I will give some counterexamples to Suppes' theory. I will then modify some of Suppes' definitions in a natural manner, and show that even with modification they are still prone to counterexamples.

The main thrust here is that Suppes' account of causation is intrinsically defective. I believe that there is no way to differentiate genuine from spurious causes or direct from indirect causes using only probability relations; thus no minor modifications of Suppes' definitions will be sufficient to resolve these difficulties. While presenting counterexamples to Suppes' definitions, I will also try to explain in principle why each particular example is a counterexample to Suppes' theory. After presenting these counterexamples, I will argue that interactive forks show that the basic intuition around which Suppes built his theory is faulty.

4.1. Exposition of Suppes' Theory

4.1.1. Prima Facie Causes

Suppes (1970, p. 12) begins his discussion of causality by giving a definition of a prima facie cause:

Definition 1: The event $B_{t'}$ is a prima facie cause of the event A_t if and only if

1. $t' < t$,
2. $P(B_{t'}) > 0$,
3. $P(A_t / B_{t'}) > P(A_t)$.

The idea here is that if the probability of one event given another event is higher than the probability of the first event alone, then the two events are causally connected in some way. Being a prima facie cause is a prerequisite for being any kind of cause other than a negative cause. Throughout this discussion Suppes assumes that the direction of causation always follows the direction of time; he does not allow for reverse causation. However, even if two events are causally connected according to definition 1, we cannot infer that the earlier is the cause of the later: they might both be effects of some common cause.

4.1.2. Spurious Causes

In order to account for this possibility, Suppes (1970, p. 23) introduces the idea of a spurious cause:

Definition 2: An event $B_{t'}$ is a spurious cause of A_t (in sense one) if and only if $B_{t'}$ is a prima facie cause of A_t and there is a $t'' < t'$ and an event $C_{t''}$ such that

1. $P(B_{t'} C_{t''}) > 0$,
2. $P(A_t / B_{t'} C_{t''}) = P(A_t / C_{t''})$,
3. $P(A_t / B_{t'} C_{t''}) \geq P(A_t / B_{t'})$.

The intuitive idea here is that a spurious cause does not change the conditional probability of the event A_t given $C_{t''}$. The addition of $B_{t'}$ into the set of factors

contributing to A_t has no real effect upon the occurrence of A_t ; event $C_{t''}$ can account for A_t at least as well as B_t can. The problem with this definition, however, is that it makes an event B_t spurious if an earlier event $C_{t''}$ satisfying the above requirements exists. Suppes himself finds this troublesome, although he neglects to explain precisely why he finds it troublesome. Perhaps the intuition he was trying to capture is that if C is a genuine cause of A and B a spurious cause of A , then the occurrence of B should be irrelevant regardless of whether C occurs. The definition of spurious cause in sense one does not reflect this intuition, which may have been the reason Suppes rejected it and developed a modified account of spuriousness.

Suppes (1970, p. 25) thinks that, if instead of demanding that an earlier event exists, we demand that a certain kind of earlier event exists, then we will have a more intuitive account of a spurious cause:

Definition 3: An event B_t is a spurious cause of A_t (in sense two) if and only if B_t is a *prima facie* cause of A_t and there is a $t'' < t'$ and a partition $\pi_{t''}$ such that for all elements $C_{t''}$ of $\pi_{t''}$

1. $P(B_t, C_{t''}) > 0$,
2. $P(A_t / B_t, C_{t''}) = P(A_t / C_{t''})$.

From now on I will abbreviate "spurious in sense two" by "spurious₂" and "spurious in sense one" by "spurious₁." This definition makes an event spurious₂ if the world can be partitioned in such a way that the above conditions are satisfied. Thus if we can observe a certain kind of event given by the partition, the observation of the later event B_t is uninformative, which makes it a spurious₂ cause. Suppes proves that if an event is a spurious₂ cause then it is a spurious₁ cause. The converse of this theorem, however, is not necessarily true: it is possible for an event to be a spurious₁ cause and not be a spurious₂ cause.

As an example of a spurious₁ cause, let us take the case of decreasing air pressure causing not only rain but also a falling barometer reading. The falling barometer reading is a *prima facie* cause of rain; given that the barometer reading is dropping, the probability that it will rain rises. Letting A denote rain, B denote a falling barometer reading, and C denote decreasing air pressure, the probability of rain given that the barometer reading and the air pressure are decreasing, $P(A/CB)$, is equal to the probability of rain given that the air pressure is decreasing,

$P(A/C)$; thus the second condition of definition 2 is satisfied. The third condition is likewise satisfied, since the probability of rain given decreasing air pressure and a falling barometer reading is at least as great as the probability of rain given a falling barometer reading, $P(A/BC) \geq P(A/B)$. Thus, by definition 2 a falling barometer reading is a spurious₁ cause of rain. The falling barometer reading is also a spurious₂ cause of rain. If we let π be our partition {decreasing air pressure, non-decreasing air pressure}, then

1. $P(BC) > 0$,
2. $P(A/BC) = P(A/C)$,
3. $P(A/B \neg C) = P(A/\neg C)$.

So the falling barometer reading is a spurious₂ cause of the rain.

4.1.3. Direct and Indirect Causes

Closely related to the notion of a spurious cause is the idea of an indirect cause. We will first define a direct cause:

Definition 4: An event $B_{t'}$ is a direct cause of A_t if and only if $B_{t'}$ is a prima facie cause of A_t and there is no t'' and no partition $\pi_{t''}$ such that for every $C_{t''}$ in $\pi_{t''}$

1. $t' < t'' < t$,
2. $P(B_{t'}, C_{t''}) > 0$,
3. $P(A_t / B_{t'}, C_{t''}) = P(A_t / C_{t''})$.

We will then define an indirect cause to be a prima facie cause that is not direct. One immediately notices the similarity between definition 3 and definition 4. The main difference is that t'' falls between t and t' in definition 4. Although Suppes does not do so, this similarity suggests that a definition of direct cause could also be developed using the analysis of a spurious₁ cause.

Definition 5: An event $B_{t'}$ is a direct cause in sense one of A_t if and only if $B_{t'}$ is a prima facie cause of A_t and for every t'' , $t' < t'' < t$, there is no $C_{t''}$ such that

1. $P(B_{t'}, C_{t''}) > 0$,

$$2. P(A_t / B_t, C_{t''}) = P(A_t / C_{t''}),$$

$$3. P(A_t / B_t, C_{t''}) \geq P(A_t / B_t).$$

The conditions of definition 5 are similar to those of definition 2 with the difference that $t' < t'' < t$. We will call the definition of direct cause given by definition 4 direct cause in sense two. From now on I will abbreviate "direct cause in sense one" by "direct₁" and "direct cause in sense two" by "direct₂." Definitions 4 and 5 say that a cause is a direct cause if and only if there is no later event (or kind of event) that will account for A_t as well as B_t does. Whereas an event is direct₂ if a certain kind of event doesn't exist, an event is direct₁ if a certain event doesn't exist. This mirrors the difference between spurious₁ and spurious₂ causes. I mentioned earlier that Suppes proved that if a cause is a spurious₂ cause, then it is also a spurious₁ cause. A similar proof could be constructed to show that if a prima facie cause is a direct₁ cause, then it is also a direct₂ cause, and if it is an indirect₂ cause, then it is an indirect₁ cause.

4.1.4. Supplementary Causes

Suppes (1970, p. 33) then enriches his theory by defining supplementary causes:

Definition 6: Events B_t and $C_{t''}$ are supplementary causes of A_t if and only if

1. B_t is a prima facie cause of A_t ,
2. $C_{t''}$ is a prima facie cause of A_t ,
3. $P(B_t, C_{t''}) > 0$,
4. $P(A_t / B_t, C_{t''}) > \max\{P(A_t / B_t), P(A_t / C_{t''})\}$.

Two causes are supplementary causes if the probability of an event occurring given both is higher than it would have been given either one alone. Thus, consuming drugs and consuming alcohol are supplementary causes of death, because the probability of dying given one has consumed drugs and alcohol is greater than either the probability of dying given one has consumed drugs or the probability of dying given one has consumed alcohol. It is worth noting that no spurious cause of

A can be a supplementary cause of A. If, according to condition (2) of definition 2 or 3, $P(A/BC)=P(A/C)$, then it is not the case that condition (4) of definition 6 can be satisfied, so B and C will not be supplementary causes.

4.1.5. Sufficient Causes

Sufficient causes are viewed as those limiting cases in which the conditional probability of an event reaches one:

Definition 7: An event B_i is a sufficient (or determining) cause of A_i if and only if B_i is a prima facie cause of A_i and $P(A_i/B_i)=1$. (1970, p. 34)

Although some philosophers deny it, the sufficient cause relation is normally assumed to be transitive; if C is a sufficient cause of B, and if B is a sufficient cause of A, then C is a sufficient cause of A. Suppes' analysis of a sufficient cause yields this result.

Theorem 8: If $P(A/B)>P(A)$, $P(B/C)>P(B)$, $P(A/B)=1$, and $P(B/C)=1$, then $P(A/C)=1$.

Proof:

- | | |
|---|---|
| 1. if $P(A/B)=1$ and $P(BC)>0$, then $P(A/BC)=1$
2. $P(A/BC)=1$
3. $P(A/C)=P(B/C)P(A/BC)+P(\neg B/C)P(A/\neg BC)$
4. $P(\neg B/C)=0$
5. $P(A/C)=1$ | Suppes'
(1970, p. 35) Theorem
1, assumptions
theorem on
total probability
from $P(B/C)=1$
2,3,4 |
|---|---|

4.1.6. Necessary Causes

Another important idea in causation is that of a necessary cause or condition. Although Suppes does not discuss them, we can define a necessary cause as follows:

Definition 9: An event B_i is a necessary cause (or condition) of A_i if and only if B_i is a prima facie cause of A_i and $P(A_i/\neg B_i)=0$.

Event B is necessary for A if and only if the probability of A given the absence of B is equal to zero. We normally think that necessary causes or conditions are also transitive, and this analysis supports that idea.

Theorem 10: If $P(A/B) > P(A)$, $P(B/C) > P(B)$, $P(A/\neg B) = 0$, and $P(B/\neg C) = 0$, then $P(A/\neg C) = 0$.

Proof:

1. If $P(A/\neg B) = 0$ and $P(\neg B/\neg C) > 0$, then $P(A/\neg B/\neg C) = 0$ theorem similar to Suppes' theorem in above proof
2. $P(A/\neg B/\neg C) = 0$ 1, assumptions
3. $P(A/\neg C) = P(B/\neg C)P(A/B/\neg C) + P(\neg B/\neg C)P(A/\neg B/\neg C)$ theorem on total probability
4. $P(A/\neg C) = 0$ 2,3,assumptions

In conjunction with the transitivity of sufficient causes, this entails that if we have a chain of necessary and sufficient causes, any member of that chain at t' is a necessary and sufficient cause of any member of that chain at t , for all $t > t'$.

4.2. Analysis of Spurious and Direct Causes

4.2.1. Example 1

I would like to assess the adequacy of these definitions with the use of several examples. Let us first consider the adequacy of the definition of a spurious cause. It seems reasonable to believe the world is composed of both deterministic and probabilistic causes; presumably if there are indeterminate events they will be intermingled with determinate events and thus there will be causal chains consisting of both deterministic and probabilistic causes. For example 1, consider the causal chain $\rightarrow D \rightarrow C \rightarrow B \rightarrow A$, where B is a probabilistic cause of A; C is a necessary and sufficient cause of B; D is a necessary and sufficient cause of C, etc. The first thing to notice is that B is a spurious cause of A since the following conditions are satisfied:

1. $P(BC) > 0$,
2. $P(A/BC) = P(A/C)$,
3. $P(A/BC) \geq P(A/B)$.

We know that condition (1) is satisfied because $P(B/C) = 1$. We can also show that condition (2) is satisfied because $P(B/C) = 1$.

Proof:

- | | |
|--------------------------------|-------------------|
| 1. $P(B/C)=1$ | assumption |
| 2. $P(B/AC)=1$ | 1,Suppes' theorem |
| 3. $P(BCA)=P(CA)$ | 2,definition |
| 4. $P(BC)=P(C)$ | 1,definition |
| 5. $P(ABC)/P(BC) = P(AC)/P(C)$ | 3,4 |
| 6. $P(A/BC)=P(A/C)$ | 5,definition |

We also know that condition (3) is satisfied because $P(B/\neg C)=0$:

Proof:

- | | |
|---------------------------|----------------|
| 1. $P(B/\neg C)=0$ | assumption |
| 2. $P(B-C)=0$ | 1,definition |
| 3. $P(AB-C)=0$ | 2 |
| 4. $P(AB)=P(ABC)+P(AB-C)$ | theorem |
| 5. $P(AB)=P(ABC)$ | 3,4 |
| 6. $P(BC)=P(B)-P(B-C)$ | theorem |
| 7. $P(BC)=P(B)$ | 2,6 |
| 8. $P(A/BC)=P(A/B)$ | 5,7,definition |

This might lead us to ask if C is a genuine cause of A. But C is also a spurious cause of A since,

1. $P(CD)>0,$
2. $P(A/CD)=P(A/D),$
3. $P(A/CD)\geq P(A/C),$

are also true. The proof of this follows from the transitivity of sufficient and necessary causes.

Proof: Since $P(B/C)=1$ and $P(C/D)=1$, we know that $P(B/D)=1$. Similarly we know that $P(B/\neg D)=0$. Thus we know that C is a spurious cause of A, by reasoning similar to the previous proof, *mutatis mutandis*.

We could continue proving that each member of such a chain is a spurious cause of A until we reached the first necessary and sufficient cause in that chain, if there is one. If there is a first necessary and sufficient cause in the chain it will be the only genuine cause of A; if the chain has no first necessary and sufficient cause, then A has no genuine cause. Both these alternatives are paradoxical. Clearly A has some genuine causes, namely, all the members of that chain.

It is also important to notice that in example 1, B is a direct₁ cause of A; there is no event falling between A and B that renders B indirect₁. Furthermore, B is the only direct cause of A; C, D, and all of the other necessary and sufficient causes in the chain are indirect₁ since

1. $P(BX) > 0$,
2. $P(A/BX) = P(A/B)$,
3. $P(A/BX) \geq P(A/X)$,

are satisfied, where X denotes any of the necessary and sufficient causes in the chain. Thus B makes all of the other members of the chain indirect₁ causes of A, and all of the other members of the chain make B a spurious₁ cause of A. In this situation we have an earlier event making a later event, B, spurious₁, and the later event, B, making the earlier event indirect₁. Clearly this is an undesirable situation. We don't want to say the only direct₁ cause of A is spurious₁ and that the only genuine cause of A (if there is one) is indirect₁. Suppes believed that spurious causes of A should not make genuine causes of A indirect causes, and he proved it could not happen for spurious₂ and indirect₂ causes. These problems arise because Suppes defines a cause to be spurious₁ if there exists an earlier event that can account for the probability of the effect just as well as any later event can. But in chains of necessary and sufficient causes, each member of a chain accounts for the probability of the last member just as well as any other member of that chain.

4.2.2. Example 2

Example 2 concerns a situation in which it is possible for an effect to be caused as a result of more than one cause. Suppose we have a very brittle glass window, a gun and a slingshot, and that someone shoots the window with the slingshot while slightly later someone else shoots the window with the gun, and the rock and bullet meet at the window and shatter it at just the same time. Let us adopt the following dictionary to symbolize this example:

W_t = the window being broken at t ,
 B_t = a bullet hitting the window at t ,
 R_t = a rock hitting the window at t ,
 $G_{t''}$ = a gun shot at the window at t ,
 $S_{t'''}$ = a slingshot shot at the window at t .

The present example might then be diagrammed as in figure 4-1.

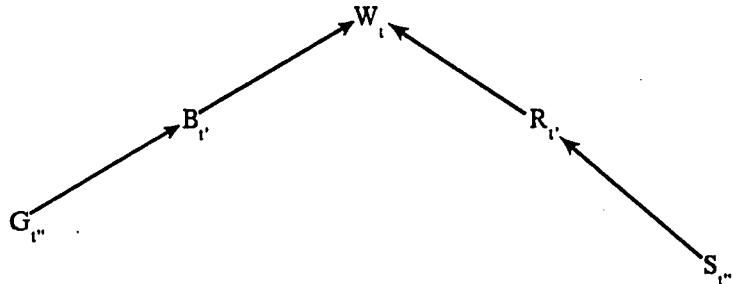


Figure 4-1: Illustration of Example 2

In figure 4-1 we have $t'' < t' < t < t'''$. However, even though B_t and R_t are genuine causes of W_t , both B_t and R_t may be spurious causes of W_t on Suppes' account. Suppose that the causes in this example are sufficient, i.e., that the following claims are true:

1. $P(W_t/B_t)=1$,
2. $P(W_t/R_t)=1$,
3. $P(B_t/G_{t''})=1$,
4. $P(R_t/S_{t''''})=1$.

In this case both the bullet hitting the window (B_t) and the rock hitting the window (R_t) are spurious causes. B_t is a spurious cause of W_t because $G_{t''}$ occurs earlier than B_t , and

1. $P(B_t/G_{t''}) > 0$,
2. $P(W_t/B_t G_{t''}) = P(W_t/G_{t''}) = 1$,
3. $P(W_t/B_t G_{t''}) \geq P(W_t/B_t) = 1$,

are true. R_t is a spurious cause of W_t because $S_{t''''}$ occurs earlier than R_t , and

1. $P(R_t S_{t''''}) > 0$,

$$2. P(W_t / R_t S_{t''}) = P(W_t / S_{t''}) = 1,$$

$$3. P(W_t / R_t S_{t''}) \geq P(W_t / R_t) = 1,$$

are true. But even more puzzling is the fact that shooting the gun at the window is a spurious cause. The slingshot being shot at the window is the only genuine cause of the window breaking because:

$$1. P(G_t S_{t''}) > 0,$$

$$2. P(W_t / G_t S_{t''}) = P(W_t / S_{t''}) = 1,$$

$$3. P(W_t / G_t S_{t''}) \geq P(W_t / G_t) = 1.$$

The only reason shooting the slingshot is picked as the genuine cause instead of shooting the gun is that shooting the slingshot occurred earlier. But this seems to be ultimately arbitrary. I see no reason to believe that the earliest cause that gives a certain probability is therefore the genuine cause.

4.2.3. Example 3

This becomes even more apparent if we change example 2 somewhat to get example 3. As before, let a slingshot be shot at the window, but in example 3 let the gun be fired at such a time that the bullet breaks the window and the rock follows immediately behind the bullet, having no impact on the glass. This example is the same as example 2 except that B_t is slightly before $R_{t+\epsilon}$; thus the rock has no effect on W_t . However, the same probability relations hold and shooting the gun is therefore a spurious cause of the window breaking while the genuine cause is shooting the slingshot, even though it was the bullet that broke the window and the rock had no causal interaction with the window. This is an example in which the "spurious" cause is really the genuine cause, and the "genuine" cause (according to definition 2) is really a spurious cause. This result is the exact opposite of what we want; it arises because Suppes treats the earliest predictively informative cause as the genuine cause, which is not always appropriate. Suppes does not account for the time it may take a causal chain to occur and achieve its effect; in this case, for instance, trouble arises because the bullet travels much faster than does the rock. It appears that mere probability relations do not take this factor into account.

This example also presents a problem for the definition of supplementary

causes given in definition 6. Suppose all of the causes are not sufficient but instead are probabilistic. In this case, shooting the gun and shooting the slingshot are supplementary causes of the window breaking. But if the bullet breaks the window, it is false to say that shooting the slingshot is a genuine supplementary cause. This supports the claim that probability relations alone cannot pick out actual causal chains.

In reply to these examples, Suppes could claim that, if we specify our events closely enough, there is a difference between a window broken by a rock and a window broken by a bullet. In that case the probability relations given above would not necessarily hold, and Suppes' definition of a spurious₁ cause might escape my criticism. I think that even if Suppes wanted to specify events that precisely, counterexamples could still be constructed. But even more importantly, I feel Suppes is committed to never specifying events that finely. Suppes wants his theory to account for our everyday use of the idea of causation, but in everyday language we do not ordinarily specify events very precisely. Thus if Suppes wishes to require that events be specified precisely in order to avoid these difficulties, he will no longer be able to claim that his theory captures our ordinary use of the concept of causation.

4.3. Analysis of Spurious₂ and Direct₂ Causes

4.3.1. Examples 2 and 3

I would now like to discuss Suppes' analysis of spurious₂ causes. Let us continue the discussion of example 3 in which a gun is fired after a slingshot is shot, yet the bullet breaks the glass and the rock has no effect on the window. The shooting of the gun was spurious₁, but it is not spurious₂. In order for it to be spurious₂ the following would have to hold:

1. $P(G_t, S_{t''}) > 0, P(G_t, \neg S_{t''}) > 0,$
2. $P(W_t / G_t, S_{t''}) = P(W_t / S_{t''}),$
3. $P(W_t / G_t, \neg S_{t''}) = P(W_t / \neg S_{t''}).$

But it should be clear that (3) is not true. If the slingshot is not shot, the bullet

shot by the gun will still break the window with a probability of 1, which is not equal to $P(W_t / \neg S_{t''})$. Thus the analysis of spurious₂ seems better off than the analysis of spurious₁, because the shooting of the gun in this example is not a spurious₂ cause. However, the shooting of the slingshot is still not considered a spurious₂ cause. But clearly the shooting of the slingshot should be considered a spurious cause in this case. Thus it appears as if the analysis of spurious₂, like that of spurious₁, looks for potential or possible causes without paying attention to actual causal chains. Furthermore, example 2 is not a counterexample to the definition of spurious₂, because if we take $\pi_{t''}$ to be the natural partition $\{S_{t''}; \neg S_{t''}\}$, then $P(W_t / G_{t''} \neg S_{t''}) \neq P(W_t / \neg S_{t''})$. If the slingshot is not shot, the shooting of the gun will cause the window to be broken.

4.3.2. Example 1

The definition of spurious₂ also handles example 1. In order for B to be a spurious₂ cause of A, $P(BC)$ and $P(B-C)$ must both be greater than zero. But if C is a necessary cause of B, $P(B-C)=0$, and thus B is not a spurious₂ cause of A. The same reasoning shows that none of the members of the chain are spurious₂ causes. Suppes (1970, p. 29) also proved a theorem showing that a partition $\{B, \neg B\}$ could not make C an indirect₂ cause, while the partition $\{C, \neg C\}$ made B a spurious₂ cause. Thus the definition of spurious₂ seems to have eluded the main thrust of example 1.

4.3.3. Example 4

Although the criterion that, $P(BC)>0$ for all members C of the partition π , saves the definition of spurious₂ from counterexamples that plague the definition of spurious₁, we shall see that this requirement brings along problems of its own. Consider example 4:

- $A_{t''}$ = Joe's wife having syphilis at time t'' ,
- $B_{t'}$ = Joe's having paresis at time t' ,
- C_t = Joe's having syphilis at time t .

In this example, let Joe's having syphilis (C) be a common cause of his wife having syphilis (A) and Joe's having paresis (B). Let us further assume that syphilis is a necessary cause of paresis and that $t < t' < t''$. Joe's having paresis is then clearly a

spurious cause of Joe's wife having syphilis, but unfortunately it is not a spurious₂ cause. In order for B to be a spurious₂ cause of A, we need a partition {C,¬C} for each member of which $P(B-C) > 0$. But since C is a necessary cause of B, $P(B-C) = 0$. Thus B is not a spurious₂ cause of A, and Joe's having paresis is a genuine cause of his wife's having syphilis, which is not the case.

This situation will arise anytime a member of the partition is a necessary cause of the event we wish to be spurious₂. It may be clearer to look at a deterministic example. Suppose we have a very good barometer in which decreasing air pressure is a necessary and sufficient cause of a falling barometer reading. But then a falling barometer reading cannot be a spurious₂ cause of a storm, because the probability of a falling barometer reading and increasing air pressure is zero. But clearly this barometer is a spurious cause of the storm. Suppes' formulation of spurious₂ causes is too strong because it excludes effects of necessary causes from being spurious₂ causes a priori. This is a serious defect in his definition of spurious₂ causes.

4.3.4. Suppes' Defense of the Requirements of Spurious₂

Suppes does attempt to defend his requirement that, for all members of the partition π , $P(BC) > 0$. He (1970, p. 36) says:

An omniscient God might object to this aspect of the definition of spurious, but for limited human knowers it seems wholly defensible.... Only a God who knows everything would have a distribution that assigns only probability one or zero to any event, and only such a distribution could never satisfy the conditions.

One might expand Suppes' comments and object that Suppes' theory is a theory of probabilistic causality and therefore does not apply to cases in which probabilities reach limiting values of zero or one. Thus my counterexamples involving causes that are necessary or sufficient would be inappropriate.

In response to these objections, I would first note that Suppes clearly intended his theory to apply to cases in which some of the probabilities are one or zero. Suppes (1970, p. 11) begins the exposition of his theory by remarking:

It should be emphasized that the deterministic concept of cause prominent in classical physics simply occupies the place of a special case in the theory to be obtained here. Roughly speaking, we obtain the deterministic theory by letting all the probabilities in question be either 1 or 0.

Suppes also defined sufficient causes, which shows that he does not consider his theory inapplicable to cases where probabilities reach these extremal values. Thus his theory ought to be able to handle my counterexamples.

I also think that Suppes' comments indicate a tension, if not an inconsistency, in his theory. Suppes tends to vacillate between claiming his theory handles cases with probabilities of one and zero and claiming that there never are any such cases for limited human knowers. As a matter of fact, it is reasonable to demand that there be some continuity between an analysis of necessary and sufficient causation and an analysis of probabilistic causation. A probabilistic theory of causation should merge with traditional theories when probabilities reach zero or one. To satisfy this requirement we shouldn't have to deny necessary or sufficient causes exist, as Suppes appears to be doing. Indeed, at other places in his monograph, Suppes (1970, p. 34) suggests that we cannot deny that the necessary and sufficient causes of classical mechanics exist and that he intends his theory to account for them:

As I emphasize throughout this monograph, the theory of causality advanced here is not meant to be tailored to the latest physics. It is designed to provide a framework for the analysis of causality in a wide variety of theories and, hopefully, in a way that will usually fit the intuitions about causality that go with a given theory.

Suppes himself would reject the idea that an analysis of causality should determine which types of causes exist and don't exist. But he wanted his theory to be applicable to the idea of causation in all theories, including classical mechanics. All of this shows that Suppes' theory is faced with the problem of causes which are either necessary or sufficient for their effects. Thus the definition of a spurious cause, which requires that for all $C \in \pi$, $P(BC) > 0$, is seen to be unjustified.

4.4. Spurious, and Direct, Causes

4.4.1. The Definition of a Spurious₃ Cause

We might find it profitable to attempt to modify the definition of spurious₂ so that it will encompass necessary causes, instead of completely rejecting the idea behind spurious₂ causes. If we ask why Suppes required that $P(BC) > 0$ for all members C of the partition π , the immediate answer is that, if $P(BC) = 0$ for some C, then $P(A/BC)$ is undefined, assuming the traditional definition of conditional probability, because dividing by zero is not allowed. The intuition Suppes was trying to capture by his definition is that the occurrence of a spurious cause should have no real effect upon the probability of the effect occurring, once we have knowledge of the genuine cause. But if this is a correct intuition, it also seems as though the non-occurrence of a spurious cause should not affect the probability of the effect, given knowledge of the genuine cause. The following theorem supports this correlation:

Theorem 11: If $P(BC) \neq 0$ and $P(\neg BC) \neq 0$, then $P(A/BC) = P(A/C)$ if and only if $P(A/\neg BC) = P(A/C)$.

Proof:

Suppose $P(BC) \neq 0$ and $P(\neg BC) \neq 0$

- | | | |
|----|---|------------|
| 1. | $P(A/BC) = P(A/C)$ | assumption |
| 2. | $P(ABC)/P(BC) = P(AC)/P(C)$ | 1 |
| 3. | $P(C)P(ABC) = P(AC)P(BC)$ | 2 |
| 4. | $P(C)P(ABC) - P(ABC)P(BC) = P(AC)P(BC) - P(ABC)P(BC)$ | 3 |
| 5. | $P(ABC)[P(C) - P(BC)] = P(BC)[P(AC) - P(ABC)]$ | 4 |
| 6. | $P(ABC)P(\neg BC) = P(BC)P(A \neg BC)$ | 5 |
| 7. | $P(ABC)/P(BC) = P(A \neg BC)/P(\neg BC)$ | 6 |
| 8. | $P(A/BC) = P(A/\neg BC)$ | 7 |

Thus if $P(BC) \neq 0$ and $P(\neg BC) \neq 0$, then $P(A/BC) = P(A/C)$ if and only if $P(A/\neg BC) = P(A/C)$.

This theorem tells us that if $P(A/BC)$ and $P(A/\neg BC)$ are defined, B is irrelevant to $P(A/C)$ if and only if $\neg B$ is irrelevant to $P(A/C)$; thus, we might want to claim that B is a spurious cause of A if and only if $\neg B$ is a spurious cause of A. With this in mind let us consider a third definition of spurious causation.

Definition 12: An event B_i is a spurious₃ cause of A_i if and only if B_i is a prima facie cause of A_i and there is a $t'' < t'$ and a partition $\pi_{t''}$ such that for all elements $C_{t''}$ of $\pi_{t''}$

1. $P(B_i, C_{t''}) > 0 \Rightarrow P(A_i / B_i, C_{t''}) = P(A_i / C_{t''})$,
2. $P(\neg B_i, C_{t''}) > 0 \Rightarrow P(A_i / \neg B_i, C_{t''}) = P(A_i / C_{t''})$.

The definition of a spurious₃ cause allows B to be a spurious₃ cause of A , even if, for some member C of the partition π , C is a necessary cause of B (i.e. $P(B-C)=0$).

4.4.2. The Definition of a Direct₃ Cause

After defining spurious₃ causes, we could now provide a definition of a direct₃ cause:

Definition 13: An event B_i is a direct₃ cause of A_i if and only if B_i is a prima facie cause of A_i and there is no t'' and no partition $\pi_{t''}$ such that for every $C_{t''}$ in $\pi_{t''}$

1. $t' < t'' < t$,
2. $P(B_i, C_{t''}) > 0 \Rightarrow P(A_i / C_{t''}, B_i) = P(A_i / C_{t''})$,
3. $P(\neg B_i, C_{t''}) > 0 \Rightarrow P(A_i / C_{t''}, \neg B_i) = P(A_i / C_{t''})$.

We will then define a prima facie cause to be indirect₃ if it is not direct₃.

4.4.3. Relations Among the Definitions of Spurious Causes

Given these definitions the following relations hold:

1. spurious₂ \Rightarrow spurious₃ \Rightarrow spurious₁
2. spurious₁ $\not\Rightarrow$ spurious₃ $\not\Rightarrow$ spurious₂
3. direct₁ \Rightarrow direct₃ \Rightarrow direct₂
4. direct₂ $\not\Rightarrow$ direct₃ $\not\Rightarrow$ direct₁
5. indirect₂ \Rightarrow indirect₃ \Rightarrow indirect₁

Proof:

1. Spurious₂ \Rightarrow spurious₃. Assume that B is a spurious₂ cause of A because of π . Then for all $C \in \pi$,

- a. $P(BC) > 0,$
- b. $P(A/BC) = P(A/C).$

Thus condition 1 of the definition of spurious₃ is satisfied. If $P(-BC)=0$ for some C, then condition 2 is satisfied for that C. If $P(-BC)>0$ for some C, then the consequent of condition 2 is also satisfied according to theorem 11. Thus spurious₂ \Rightarrow spurious₃.

2. Spurious₃ \Rightarrow spurious₁. Suppose B is a spurious₃ cause of A because of {C, \neg C}. Then we have 4 possible cases to look at.

- a. $P(BC) > 0$ and $P(B-C) > 0.$ If this is true then B is also a spurious₂ cause of A. But Suppes has proven that all spurious₂ causes are spurious₁ causes. Thus B is a spurious₁ cause.
- b. $P(BC) > 0$ and $P(B-C) = 0.$ Since B is a spurious₃ cause, $P(A/BC) = P(A/C)$, so condition 2 of definition 2 is satisfied. $P(A/BC) = P(ABC)/P(BC).$ Also, $P(BC) = P(B) - P(B-C) = P(B)$, since $P(B-C) = 0.$ $P(ABC) = P(AB) - P(AB-C) = P(AB)$, since $P(B-C) = 0.$ So $P(ABC)/P(BC) = P(AB)/P(B).$ Thus $P(A/BC) = P(A/B)$, and condition 3 is satisfied.
- c. $P(BC) = 0$ and $P(B-C) > 0.$ The proof of this is the same as for the previous case except we show that $P(A/B-C) = P(A/\neg C)$ and $P(A/B-C) = P(A/B).$
- d. $P(BC) = 0$ and $P(B-C) = 0.$ But $P(B) = P(BC) + P(B-C) = 0$, which is false, since B is a prima facie cause of A. So this is not a legitimate case.

Thus spurious₃ \Rightarrow spurious₁.

3. Spurious₁ $\not\Rightarrow$ spurious₃. Let A=a disease; C=an injection of bacteria known to cause A; and B=being exposed to people who have this disease. Suppose $P(A/C) = 1$, $P(A/\neg C)$ is very low, and $P(A//B) = 1/2.$ Then $P(A/BC) = P(A/C) \geq P(A/B)$, so B is a spurious₁ cause of A. But B is not a spurious₃ cause, since $P(A/B-C) \neq P(A/\neg C).$ So being a spurious₁ cause does not imply being a spurious₃ cause.

4. Spurious₃ $\not\Rightarrow$ spurious₂. In example 4, Joe's having paresis was not a spurious₃ cause of his wife's having syphilis, but it is a spurious₃ cause. Thus some spurious₃ causes are not spurious₂ causes.

Since the definitions of direct and indirect causes mirror the definitions of spurious causes, the rest of the table can be proved *mutatis mutandis*.

This shows that spurious_3 is a stronger condition than spurious_1 , but is a weaker condition than spurious_2 .

4.4.4. Spurious₃ Causes and Examples 2, 3, and 4

The definition of spurious_3 handles examples 2 and 3 in the same manner spurious_2 does, so I won't repeat that discussion. But spurious_3 handles example 4 much better than spurious_2 did. Joe's having paresis is a spurious_3 cause of Joe's wife having syphilis, because the following is true:

$$\begin{aligned} P(A_{\text{r}}/B_{\text{r}}C_{\text{r}}) &= P(A_{\text{r}}/C_{\text{r}}), \\ P(A_{\text{r}}/\neg B_{\text{r}}\neg C_{\text{r}}) &= P(A_{\text{r}}/\neg C_{\text{r}}). \end{aligned}$$

Thus the definition of spurious_3 handles example 4 the way it should and affords a more satisfactory definition than does that of spurious_2 .

4.4.5. Spurious₃ Causes and Example 1

The requirement that for all members C of the partition π , $P(BC)>0$, enabled the definition of spurious_2 to escape the counterexample of example 1. The definition of spurious_3 relaxes that requirement somewhat and therefore cannot handle example 1 satisfactorily. The only genuine cause according to the definition of spurious_3 will be the first necessary and sufficient cause in the chain, if there is one; if the chain stretches infinitely into the past, A will have no genuine cause. This can be proven as follows. Consider the following theorem:

Theorem 14: If $P(B/\neg C)=0$, then $P(\neg B/\neg C)=1$

Now consider any member of the necessary and sufficient causal chain, say G. If I is a member of the chain before G, the following holds:

1. $P(GI)>0$, $P(\neg GI)=0$, $P(G\neg I)=0$, $P(\neg G\neg I)>0$,
2. $P(A/GI)=P(A/I)$,
3. $P(A/\neg G\neg I)=P(A/\neg I)$.

That (1) holds is essentially a result of the above theorem, and the transitivity of necessary and sufficient causes. (2) and (3) follow from our discussion of example 1 in section 4.2.1. If there is no member of the chain before G, G will be the

first necessary and sufficient cause in the chain and the only genuine cause of A. If there is no first cause, then there is no genuine cause of A. This is the same problem confronted by the definition of spurious₁.

The definition of indirect₃ also makes B the only direct cause of A, while all of the other members of the chain are indirect₃ causes. What happens is the partition {B,¬B} makes C, D, etc. indirect₃, while the partition {C,¬C} makes B a spurious₃ cause. The partition {B,¬B} makes a previous member of the chain, say X, indirect₃, since the following hold:

$$\begin{aligned} P(A/BX) &= P(A/B), \\ P(A/\neg B-X) &= P(A/\neg B). \end{aligned}$$

The proof that these hold is essentially the same as that given in section 4.4.1. Again, this situation is clearly undesirable. The definition of spurious₃ is no better off than the definition of spurious₁ in regard to example 1. The definition of spurious₂ was able to handle example 1 because it excluded many spurious causes from the class of spurious₂ causes. In the definition of spurious₃, we attempted to bring those spurious causes into the class of spurious₃ causes, but in doing so we brought along many genuine causes.

The basic problem facing a definition of spurious causation which depends solely on probability relations among events seems to be that there is no way to distinguish between a causal chain and a fork when the probabilities involved are the same. Suppose we have the chain of necessary and sufficient causes C→B→A and a causal fork where C is a necessary and sufficient common cause of A and B. All of the probability relations in these two cases will be identical, yet in one case B is a genuine cause of A and in another case B is a spurious cause of A. Thus there appears to be no way to distinguish genuine from spurious causes using only probability relations among the events, which is a very serious problem for a probabilistic theory of causality.

4.5. Spurious Causes and Interactive Forks

Another serious problem for Suppes' definitions is the existence of interactive forks. Earlier we saw that in an interactive fork ACB, where C is the common interaction, that $P(A/BC) > P(A/C)$. This condition guarantees in the interactive fork ACB that B is not a spurious cause of A in any of our three senses. The event B is predictively informative as to the occurrence of A; thus it is a genuine cause of A as well as often being a direct cause of A in one of the three senses, according to Suppes' theory. But it should be obvious that one prong of the fork is not really a genuine cause of the other prong and that Suppes' analysis errs at this point.

Interactive forks are problematic for Suppes' theory of causation because the state of one of the prongs is the best way to predict the state of the other prong: the common cause of both prongs is less informative than knowledge of either of the prongs. But the intuition around which Suppes built his theory is that only genuine causes are predictively informative, and that spurious causes are predictively uninformative. Interactive forks are not accounted for by Suppes' intuitions; thus the more faithfully Suppes' theory reflects his intuition, the more serious will be the problem posed by interactive forks. Suppes has oversimplified by assuming that all genuine causes add predictive power and that all spurious causes are predictively uninformative.

4.6. Final Remarks on Suppes

In this chapter I have attempted to present the basic features of Suppes' theory of probabilistic causality and to critically appraise the adequacy of that theory. I initially questioned his definitions of spurious causes. Since some causal chains have exactly the same probability relations as causal forks, it is impossible to distinguish between them using only probability relations. And it is then impossible to distinguish genuine causes from spurious causes using only probability relations. I have also questioned Suppes' basic intuition that spurious causes are predictively uninformative, once knowledge of a genuine cause is known. This is especially evident when interactive forks are considered. For interactive forks we found that

either prong may be the best predictor of the state of the other, even though it is only a spurious cause of that prong. This shows that Suppes' basic intuition is at fault, since even if his theory were to capture his intuitions, it would still be unable to handle interactive forks. Suppes also defended the idea that causes are always positively relevant to their effects; we shall return to Suppes' defense of this claim in a later chapter.

CHAPTER 5

GOOD'S QUANTITATIVE CAUSAL CALCULUS

A quite different approach to probabilistic causality is given to us by I. J. Good. Good's theory is different from the other theories that we have discussed so far in that his theory is a quantitative theory; Good claims to explicate the strength of causal chains and nets. No other author on probabilistic causality makes this claim. At first glance one may be tempted to think that Good has succeeded, because his paper is very formal and mathematical. Good presents 24 axioms, 18 numbered theorems, and numerous other definitions and theorems. Unfortunately there is not a sufficient amount of discussion of the axioms and theorems to render them clear to the reader. Several of his axioms have no explanation along with them, and his notation appears to be continually changing. One can easily understand why Good's articles have been ignored by the philosophical community.

5.1. Two Different Causal Concepts

The basis of Good's theory of probabilistic causality lies in the explication of two related, but supposedly quite different causal concepts. Good distinguishes between the tendency of F to cause E or the causal support for E provided by F, from the degree to which F caused E or the contribution to the causation of E provided by F. Good expresses the tendency of F to cause E by a function $Q(E:F)$, and he expresses the degree to which F caused E by a function $\chi(E:F)$. The difference between Q and χ is not obvious; the best explication of the difference between them comes from an example that Good placed in an appendix. Good's (1961, p. 318) example is:

Sherlock Holmes is at the foot of a cliff. At the top of the cliff, directly overhead, are Dr. Watson, Professor Moriarty, and a loose boulder. Watson, knowing Moriarty's intentions, realises that the best chance of saving Holmes's life is to push the boulder over the edge of the cliff, doing his best to give it enough horizontal momentum to miss

Holmes. If he does not push the boulder, Moriarty will do so in such a way that it will be nearly certain to kill Holmes. Watson then makes the decision (event F) to push the boulder, but his skill fails him and the boulder falls on Holmes and kills him (event E).

This example shows that $Q(E:F)$ and $\chi(E:F)$ cannot be identified, since F had a tendency to prevent E and yet caused it. We say that F was a cause of E because there was a chain of events connecting F to E, each of which was strongly caused by the preceding one.

From this example it is difficult to see why Q and χ cannot be the same. It is easy enough to see that F caused E, and to determine the meaning of χ , but one wonders what the meaning of Q is. From this example it almost appears as if the tendency to prevent E is merely an intention to prevent E. However, we can be certain that this is not what Good had in mind when he was describing Q. Another possibility is that χ deals with actual causal chains and Q deals with general causal laws, but this too is untenable, given the rest of Good's discussion. The unclarity as to the distinction between χ and Q is best resolved by looking at the axioms and definitions instead of looking at the example Good gives.

5.1.1. The meaning of $\chi(E:F)$

The various axioms and theorems make it clear that χ is a measure of the strength of the causal net connecting two events. Good's theory is general enough to account for events related in complicated nets, and not merely in causal chains. The meaning of χ is given by axiom 9:

- A9 $\chi(E:F)$ is the strength of the complete causal net joining F to E. More precisely, it is the limit, as the sizes of the events tend uniformly to zero, of the strengths of nets; where each net of the sequence joins F to E, consists of a finite number of events, and omits no events temporarily between F and E.(1961, p. 311)

The function χ is a measure of the strength of the whole net connecting two events, including all intermediate events.

5.1.2. The Meaning of $Q(E:F)$

The meaning of the function Q is harder to discover. Good claims in axiom 10 that if a causal net consists only of events F and E , then $\chi(E:F)=Q(E:F)$, if $Q(E:F)$ is positive, and otherwise $\chi(E:F)=0$. Good also remarks that events later than F and earlier than E may affect the value of χ , but not of Q . The function $Q(E:F)$ only depends on three probability relations: $P(E/F)$, $P(E/\neg F)$, and $P(E)$. From this we can see that Q will be some sort of measure of statistical relevance. Good (1961, p. 317) finally presents his definition of Q in theorem 19:

$$Q(E:F)/G = \log[P(\neg E/\neg F \& G)/P(\neg E/F \& G)]$$

The G is the background information that is conditionalized on. Although Good usually leaves off the G , we must remember that it is there. Good (1961, pp. 308-9) requires that both Q and χ must always be conditional on all true laws of nature, whether known or unknown, and the essential physical circumstances just before F started. We will denote all true laws of nature by H , and the essential physical circumstances just before F started by U . The assumptions of H and U will be important at a later stage of our discussion.

We thus see that $Q(E:F)$ is really an abbreviation for $Q(E:F/U\&H)$, since we are always conditionalizing on all true laws of nature and the essential physical circumstances just before F started. But even $Q(E:F/U\&H)$ is not precise notation. Good thinks that the expression " F caused E " is an abbreviation for " F , as against $\neg F_D$, caused E , rather than E' ." Thus our more complete notation would look like this: $Q(E:F|F_D/U\&H\&(EvE'))$. We will normally not use this complicated notation, but we should remember that $Q(E:F)$ is an abbreviation.

5.1.3. $\chi(E:F)$ and $Q(E:F)$

At this point let us return to Good's example of Holmes and Moriarty. Watson's decision to push the boulder had a tendency to prevent Holmes' death; this means that Watson's decision had positive statistical relevance to preventing Holmes' death. But we also see that there was a strong causal chain connecting Watson's decision with Holmes' death. Causal tendencies are measures of statistical relevance, and the strength of actual causal chains are measures of the strength of the individual links in the chain. We saw above that if we have a chain consisting of

only F and E, then Q and χ have the same value. This means that the strength of a causal net will be a function of the tendency of each of the links to cause the next link. Thus χ is a function of Q.

5.2. Good and Spurious Correlations

5.2.1. Salmon's Objection

The only philosopher to publish any discussion of Good's theory that I know of is Salmon. Salmon correctly notices that the function Q is simply a measure of statistical relevance. But Salmon (1980, p. 52) points out that a statistical correlation does not necessarily mean that one event causes another, or even has a tendency to cause the other event. This is easily seen in the famous example of the falling barometer reading being correlated with storms. Even though the falling barometer reading is statistically correlated with the storm, nobody seriously considers that the falling barometer reading has a tendency to cause a storm.

5.2.2. Good's reply to Salmon

Shortly after Salmon's criticism's of Good were published, Good furnished us with a reply. Good replies to the above criticisms by agreeing that Q is a measure of statistical relevance, but that it is a measure of statistical relevance that is conditional on H and U. Good (1980, p. 302) claims:

But when U (and H) are given, or as some statisticians would say are "partialed out", the falling reading is not statistically relevant to the storm after all. Thus this example does not begin to refute my theory.

Good's reply was the obvious fact that there are some factors that will screen off the falling barometer reading from the occurrence of a storm. Evidently Good believes that by making Q conditional on U and H this problem will be avoided. At first glance that appears plausible, but a deeper investigation is warranted.

5.2.3. Discussion of Good's reply

Let us return to Good's definition of $Q(E:F/G)$ and see if conditionalizing on H really solves the problem. Letting A designate the storm, and B designate the falling barometer reading, the definition of $Q(E:F/G)$ gives us the following:

$$Q(A:B/H) = \log[P(\neg A/\neg B \& H)/P(\neg A/B \& H)]$$

For purposes of illustration, let us assume that our barometer is a very reliable indicator of storms: thus $P(\neg A/\neg B \& H)$ is close to one, and $P(\neg A/B \& H)$ is close to zero. In this case $Q(A:B/H)$ will be the log of a very large number. Thus $Q(A:B/H)$ will be quite large, and we are justified in believing that the falling barometer reading has a strong tendency to cause the storm. Conditionalizing on H , all true causal laws, has no real effect upon whether the falling barometer reading has a strong tendency to cause a storm. Once one realizes this it is fairly easy to look at the definition of $Q(A:B/H)$ and see that the addition of H is not going to help the problem of spurious correlations. There is nothing in the definition of $Q(A:B/H)$ to separate genuine from spurious correlations. If someone asks for the probability that a storm will not occur given that the barometer reading is not falling, we normally automatically conditionalize on the true laws of nature that we know of; it would be unusual to find someone who attempted to not conditionalize on the known laws of nature. Thus the additional requirement that we conditionalize on all true laws of nature is seen to have no effect on Q .

Another way of seeing this point is to realize that conditionalizing on all true laws of nature in no way affects the probability of a storm not occurring given that the barometer reading is not falling, because it is also a true law of nature that there is a statistical correlation between falling barometer readings and storms. Consider the following principle:

$$P[X/Y \ \& \ P(X/Y)=n]=n. \quad (5.1)$$

Principle (5.1) certainly seems to be true, if n is greater than zero and less than one. If we know that the probability of X given Y is n , and that we have a Y , then it certainly seems reasonable that the probability of a X is n . Now suppose that we know that Y is not a cause of X , that $P(\neg X/\neg Y)=n$, and whatever other true laws of nature one wishes. I claim then that $P(\neg X/\neg Y \& H)=n=P(\neg X/\neg Y)$. Conditionalizing on all true laws of nature has no effect on the above probability,

which is a part of Q. Similar reasoning will show that conditionalizing on all true laws of nature will have no effect on Q.

One might object to the above reasoning by claiming that we are conditionalizing on other laws of nature besides the one that $P(X/Y)=n$. This is true, but I believe that factor will also not affect the probability of $P(X/Y)$. Suppose we consider principle (5.2):

$$P[X/Y \text{ & all true laws of nature} \text{ &} P(X/Y)=n]=n. \quad (5.2)$$

Principle (5.2) seems to be true, if $0 \leq n \leq 1$. But (5.2) is exactly the principle that we need to show that conditionalizing on H does not affect the value of Q. It is not possible to claim that making $Q(E:F)$ conditional on all true laws of nature results in the fact that one effect of a common cause does not have a strong tendency to cause the other effect of the common cause.

Let us now investigate whether making $Q(E:F)$ conditional on the essential physical circumstances just before F started, U, enables Good to distinguish spurious from genuine causes. The definition of Q tells us that

$$Q(A:B/U)=\log[P(\neg A/\neg B \& U)/P(\neg A/B \& U)]$$

In this example, U contains the information that the air pressure is dropping. We should notice that the falling air pressure screens off a falling barometer reading from the storm; thus we see that $P(\neg A/\neg B \& U)=P(\neg A/B \& U)$. This equality holds because U screens off both B and $\neg B$ from the occurrence of a storm; thus adding B or $\neg B$ to the information conditionalized on makes no difference to the probabilities. We can now see that $Q(A:B/U)=\log(1)=0$. This is what Good wants; the strength of the causal chain is zero. Thus Good has managed to avoid the problem of spurious causes that Salmon raised by conditionalizing on the essential physical circumstances just before B occurred.

However, a similar problem for Good arises if we consider interactive forks. Suppose we have a situation such as in the Einstein, Podolsky, and Rosen experiment. If we have two electrons which are in a single system, and separate them into two separate systems, it is a law of nature that spin is conserved in the systems. Thus if one of the particles has spin up in a certain direction, the other one will have spin down in that direction. In this example, we let B denote the measurement of spin up on one of the particles, and A be the other particle having

spin down in the same direction. Even by conditionalizing on U in this example, Q will be very large, which is not what we would like. We generally think that there is no causal connection between the two particles once they are separated into separate systems, and thus the value of $Q(A:B)$ should be zero. Conditionalizing on U and H will solve many of the problems of spurious correlations, but it will not solve all of them.

5.2.4. Partially Spurious Correlations

One further problem lies in Good's treatment of correlations that are *partially spurious*. Good believes that his explication helps to analyze such problems. Consider the following example given by Good (1961, p. 44):

Smoke and dust might be a strong cause of lung cancer, but smoking only a weak cause. Even so, the correlation between smoking and lung cancer may be high if there is more smoking per head in smoky districts.... Note that

$$Q(E:F \& G | \neg F \& \neg G) = Q(E:G / \neg F) + Q(E:F / G),$$

so that the tendency to cause can be split into components, somewhat in the manner of an analysis of variance. For example, the tendency for lung cancer to be caused by smoking and living in a smoky district as against not smoking and living in a clean district is equal to the tendency through living in a smoky district, given no smoking, plus the tendency through smoking, given that the district is smoky. It is also equal to the causal tendency through living in a smoky district, given that one smokes, plus the tendency through smoking, given that the district is clean.

I find this principle implausible. Suppose we change the example slightly and let E designate death, F designate the consumption of poisonous acid, and G designate the consumption of poisonous alkali. Consuming either acid or alkali alone is a sure cause of death, but together they neutralize each other and have no effect. In this case, $Q(E:F \& G | \neg F \& \neg G)$ will be low, but $Q(E:G / \neg F)$ will be high. Thus it is impossible for the above equality to be satisfied. It seems as if Good's method of dealing with partially spurious correlations points to a real flaw in his theory.

However a close investigation of Good's method of dealing with partially spurious causes indicates that Good made a mistake in his claim that

$$Q(E:F \& G | \neg F \& \neg G) = Q(E:G / \neg F) + Q(E:F / G) \quad (5.3)$$

was true. John Pollock has noted that this equation is false. Let p denote $P(\neg E/\neg(F \& G))$, q denote $P(\neg E/F \& G)$, r denote $P(\neg E/\neg G \& \neg F)$, and s denote $P(\neg E/G \& \neg F)$. Then by Good's definition of Q , we get

$$Q(E:F \& G | \neg F \& \neg G) = \log[p/q] \quad (5.4)$$

and

$$Q(E:G/\neg F) + Q(E:F/G) = \log(r/s) + \log(s/q)$$

But simple properties of logarithms give us

$$Q(E:G/\neg F) + Q(E:F/G) = \log[(r/s)*(s/q)],$$

which is equivalent to

$$Q(E:G/\neg F) + Q(E:F/G) = \log(r/q). \quad (5.5)$$

When we replace equations (5.4) and (5.5) into equation (5.3) we find that

$$\log(p/q) = \log(r/q). \quad (5.6)$$

But equation (5.6) is not true. In order for (5.6) to be true, we would have to have $p=r$, or

$$P(\neg E/\neg(F \& G)) = P(\neg E/\neg F \& \neg G).$$

But this is generally false. It appears as if Good misused DeMorgan's theorem in his derivation of equation (5.3). Thus he really has no way of dealing with partially spurious causes.

The above discussion shows that Good was not able to successfully use probability relations to distinguish genuine causal tendencies from spurious correlations. Since he was unable to do that for simple chains consisting of two events, it is obvious that the problem still remains for causal nets. Since the explication of χ depends upon the strength of the individual links, which is given by the function Q , we can see that if Q treats a spurious correlation as a genuine causal tendency, that the function χ will also treat it as a genuine causal factor. The problem is only compounded, and not solved, when we look at more complicated nets.

5.3. The Strength of Causal Chains

Although Good's theory is inadequate to distinguish spurious correlations from genuine causal tendencies, we must not forget the rest of his theory. Good has undertaken the monumental task of attempting to give a quantitative account of causation, and we should investigate how successful this aspect of his work is independently of the issue of spurious correlations. A major assumption of Good's theory, given in axiom 11, is that the strength of a causal chain is a function of the strength of the individual links in the chain. Several other axioms elaborate other aspects of this assumption. This axiom looks very plausible at first, perhaps because we believe that the strength of an ordinary chain that one buys at a hardware store is a function of the strength of its links. But this analogy may be misleading. Salmon presents a counterexample to show that the strength of a chain cannot be a function of the strength of the individual links. The basic structure of his example is simple, whereas the actual details are more complicated. Basically, Salmon describes a situation in which we have two chains, consisting of three events each. In these two simple chains, the strengths of the first two links are the same and the strengths of the two final links are the same. Thus if the strength of the whole chain is a function of the strengths of the individual links, the strengths of the two chains should be equal. But in the situation Salmon describes it seems obvious that the strengths of the two chains are different, even though the strengths of their individual links are the same.

5.3.1. Salmon's Counterexample

Salmon's (1980, pp. 52-3) counterexample involves a simple game consisting of a tetrahedron and some cards:

Consider the following simple two stage game. The player first tosses a fair tetrahedron with sides marked 1,2,3, and 4. If the tetrahedron comes to rest on any side other than 4--i.e., if side 4 shows after the toss--the player draws from a deck which contains 16 cards, 12 of which are red and 4 of which are black; if side 4 does not show, he draws from a deck containing 4 red and 12 black. On a given play, the player tosses the tetrahedron (event F) and it comes to rest with side 4 showing, so that the player draws from the first above-mentioned special deck, with the result that he gets a red card (event E).

Assuming that events F, G, and E actually occur, we should inquire into the degree that F caused E, or $Q(E:F)$. Assuming that the only way to get to draw from the special deck is to play the game, Salmon uses the above probability values to determine that $P(E/F)=10/16$.

Salmon (1980, p. 53) then asks us to consider a second game, that is very similar to the first:

...let us consider another game which is just like the foregoing except that different decks of cards are used. The player tosses a fair tetrahedron (event F') and if side 4 shows he draws from a deck containing 14 red cards and 2 black cards (event G'). If the tetrahedron comes to rest on side 4, the player draws from a deck containing 10 red cards and 6 black cards (event $\neg G'$). In this game the probability of drawing a red card (event E') equals, 13/16.

Salmon has described two causal chains, $F \rightarrow G \rightarrow E$ and $F' \rightarrow G' \rightarrow E'$, and he calculates the strengths of the links and chains. Since $P(\neg E/\neg G)=3/4$ and $P(\neg E/G)=1/4$, we can use Good's definition of $Q(E:F)$ to calculate that $Q(E:G)=\log 3$. Similarly, since $P(\neg E'/\neg G')=3/8$ and $P(\neg E'/G')=1/8$, we find that $Q(E':G')=\log 3$. This tells us that the strengths of the two links, $G \rightarrow E$ and $G' \rightarrow E'$, are equal. Salmon then notices that $Q(G:F)=Q(G':F')$, since $P(G/F)=P(G'/F')$ and $P(G/\neg F)=P(G'/\neg F')$. Thus the strengths of the two links $F \rightarrow G$ and $F' \rightarrow G'$ are equal also. Salmon (1980, p. 54) has shown that the strengths of the corresponding two chains are equal.

Salmon must now show that even though Good's theory claims the strengths of the two chains are equal, that we intuitively think that they are not of equal strength. The first thing to notice is that the probabilities of the last members of the chains given the first members of the chains are different: $P(E/F) \neq P(E'/F')$. The next thing to notice is that the statistical relevance of the first members to the last members is different: $P(E/F)-P(E/\neg F) \neq P(E'/F')-P(E'/\neg F')$. These two facts lead Salmon to believe that the strengths of the two causal chains are different.

The above example may seem somewhat contrived, but the basic features of it can be found in less contrived situations. Salmon (1980, p. 54) presents an example with exactly the same structure, but which seems more concrete:

Suppose that two individuals, Joe Doakes and Jane Bloggs, suffer from sexual disabilities. Joe is impotent and Jane is frigid. Each of them decides to seek psychotherapy. There are two alternative types of therapy available, directive or non-directive. When Joe seeks out a

psychotherapist (event F), there is a probability of $3/4$ that he will select a directive therapist and undergo that type of treatment (event G), and a probability of $1/4$ that he will select a non-directive therapist and undergo that type of treatment (event G'). If he is treated by a directive therapist, there is a probability of $3/4$ that he will be cured (event E), and if he is treated by a non-directive therapist, there is a probability of $1/4$ that he will be cured....

When Jane seeks out a psychotherapist (event F'), there is a probability of $3/4$ that she will select a directive therapist (event G'), and a probability of $1/4$ that she will select a non-directive therapist (event $\neg G'$). If she is treated by a directive therapist, there is a probability of $7/8$ that she will be cured (event E'), and if she is treated by a non-directive therapist, the probability of a cure is $5/8$.

This example has exactly the same structure as the previous example with the tetrahedron and the cards. Salmon claims that not enough information has been given to decide if the two chains have the same strength. The strength of the chains will depend upon the probability of a cure for their problems given that they do not undergo psychotherapy. For example, suppose that there is a high probability, around $3/4$, of frigidity being cured without psychotherapy, while there is a low probability, around $1/100$, of impotence being cured without psychotherapy. Concerning this, Salmon (1980, p. 55) says:

...it seems intuitively clear that the causal contribution to the cure in the case of Joe is much greater than it is in the case of Jane. For Jane's problem, the probability of cure if she undergoes therapy ($13/16$) is only slightly higher than the probability of spontaneous remission ($12/16$), but for Joe's problem, the probability of a cure if he undergoes psychotherapy ($10/16$) is much greater than the probability of spontaneous remission ($1/100$).

Salmon is claiming that in order to determine the causal contribution of a cause to its effect we have to know the likelihood of the effect occurring when the cause is not present. An actual probability value does not tell us something about the contribution of a cause to an effect. In this example it does seem reasonable that psychotherapy contributes more to the cure of Joe's problem than it does to the cure of Jane's problem. Thus Salmon concludes that the strength of a causal chain cannot be a function of the strengths of the individual links.

5.3.2. Good's reply

Good (1980, pp. 302-3) replies to Salmon's counterexample by claiming that he does not see why the probability of spontaneous remission has anything to do with the strength of the causal chain:

In this example he considers a Markov chain $F \rightarrow G \rightarrow E$ in which all three events actually occur. He implies that the strength of the chain should depend not just on $P(E|G)$, $P(E|\neg G)$, $P(G|F)$, and $P(G|\neg F)$, as in my "causal calculus", but that it should also depend at least on $P(E|\neg F)$. Since we are dealing with a Markov chain I do not agree with Salmon's intuition on this point, and I do not see in what way his examples support his case except that what is said three times sounds true.... Of course $Q(E:F)$ does depend on $P(E|\neg F)$, and until you have trained yourself to distinguish sharply between Q and χ , ...you might have a residual feeling that the strength of the Markov chain should depend on $P(E|\neg F)$ as well as the four other probabilities mentioned.

Evidently Good believes that the strength of the chain, the degree to which F caused E , or the contribution to the causation of E provided by F , is independent of $P(E|\neg F)$. recalling Salmon's example with Joe and Jane, it certainly does seem that therapy contributed to Joe's cure more than it contributed to Jane's cure. Thus if χ is to capture the meaning of the contribution to the causation of E provided by F , we must claim that Good's theory is inadequate.

Good could easily reply that he is interested in the strength of the chain from F to E , which is different from the contribution to the causation of E provided by F . The problem with this is that Good (1961, p. 307) does say that they are the same. If Good means something different from the contribution to E provided by F when he is discussing the strength of the chain from F to E , he certainly has not told us what he means. I feel that until we are given further explanation of what Good means by the strength of a chain or the contribution of F to the causation of E , we must claim that Good was incorrect in making the strength of the chain a function of the strengths of the links.

Salmon's third criticism of Good's theory concerns positive statistical relevance. Often there will be events in a chain that are not positively relevant to the next event in the chain. But Good requires that each event in the chain be positively relevant to the next event in the chain, which would exclude several causal chains from being chains according to Good's theory. We will discuss Good's answer to this objection in a later chapter.

5.4. Quantitative Theories of Causality

Good attempted to give a quantitative account of causation, which is a very difficult task. It seems as if one should concentrate first upon being able to adequately pick out probabilistic causes and effects, and then to attempt to characterize the strengths of various causal chains. Although Good's theory fails, it is instructive to see why it fails. We shall later return to Good when we discuss the problem of positive statistical relevance.

CHAPTER 6

THE QUESTION OF POSITIVE RELEVANCE

All of the various theories of probabilistic causality that we have examined so far require that a cause raise the probability of its effect. Reichenbach's definitions of causal relevance, causal betweenness, and conjunctive forks all have this property. Similarly, Suppes required that in order to be a *prima facie* cause, an event must raise the probability of its effect. If a cause lowers the probability of its effect it is a negative cause. Good also required that causes be positively relevant to their effects. But even though this idea of positive statistical relevance lies at the heart of all of these theories of probabilistic causality, there is really no defense of the doctrine. Recently, the doctrine that a cause must raise the probability of its effect has come under attack. Germund Hesslow has presented an example which he considers to be a *reductio ad absurdum* against the very idea of probabilistic causality. By attacking the central idea of probabilistic causality, that causes raise the probability of their effects, Hesslow has tried to show the futility of any theory of probabilistic causality. The different responses to the attacks on positive relevance are worth studying.

6.1. Prima Facie Causes and Negative Relevance

6.1.1. Hesslow's Argument

To begin our discussion, let us consider an example, proposed by Hesslow (1976, p. 291), which seems to suggest that a cause may lower the probability of its effect:

The basic idea in Suppes' theory is of course that a cause raises the probability of its effect, and it is difficult to see how the theory could be modified without upholding this thesis. It is possible however that examples could be found of causes that lower the probability of their effects. Such a situation could come about if a cause could lower the probability of other more efficient causes. It has been claimed, e.g., that

contraceptive pills (C) can cause thrombosis (T), and that consequently there are cases where C caused T. But pregnancy can also cause thrombosis, and C lowers the probability of pregnancy. I do not know the values of $P(T)$ and $P(T/C)$ but it seems possible that $P(T/C) < P(T)$, and in a population which lacked other contraceptives this would appear a likely situation. Be that as it may, the point remains: *it is entirely possible that a cause should lower the probability of its effect.*

I agree with Hesslow that it appears possible that $P(T/C) < P(T)$. But that is only because the taking of contraceptives lowers the probability of pregnancy. Thus we might note that $P(T/C \& \neg P) > P(T/\neg P)$ might be true, where P denotes being pregnant. Suppes (1970, p. 42) has noted that his definition of a prima facie cause can easily be relativized to background information. Thus we might change his definition of a prima facie cause to:

Definition 1: B_i is a prima facie cause of A_i with respect to information C_i , if and only if

1. $t' < t$,
2. $P(B_i \& C_i) > 0$,
3. $P(A_i / B_i \& C_i) > P(A_i / C_i)$.

In Hesslow's example we can let $\neg P$ be the background information. Then it follows that with respect to that background information the taking of contraceptives (C) is a prima facie cause of thrombosis (T).

6.1.2. Probabilistic Analogue of a Causal Field

I think that this idea of Suppes becomes much more plausible if we investigate the traditional notion of a causal field. Some philosophers (Mackie 1974) have claimed that all causal statements are made in reference to some background information or assumptions called a causal field. It is really incorrect to say that A caused B; what should be said is that in a certain causal field F, A causes B. An event A may be a cause of B in a field F, and not be a cause of B in a field G. The natural counterpart of a causal field in probabilistic causality is the reference class, or what is conditionalized on. An event may raise the probability of another event in one reference class, and lower it in another. Thus, Suppes

could reply to Hesslow by saying that, in certain reference classes or causal fields the taking of contraceptives is a *prima facie* cause of thrombosis, while in other reference classes it is not a *prima facie* cause. The question then becomes, what is the appropriate reference class? Suppes' answer to this is somewhat surprising.

6.2. Rosen's Defense of Positive Relevance

6.2.1. Rosen and Adequate Conceptual Frameworks

Deborah Rosen, a former student of Suppes, responds to Hesslow in two slightly different ways. Both of Rosen's answers deal with the conceptual framework that the causal statement is relative to. Rosen points out that Suppes has a strong belief that causal relationships are always relative to some conceptual framework. Evidently Suppes believes that what counts as a cause of another event changes as we look at different conceptual frameworks. Rosen (1978, p. 606) does not elaborate on this idea of causation, but she does point out that Suppes thinks that there are three basic frameworks in which we have causal relations: a particular scientific theory, an experiment or set of experiments, and the framework of general beliefs that expresses all the information available to us. It is not clear what it means for causation to be relative to a conceptual framework; certainly we can change probability relations by choosing different information to conditionalize on. But Suppes claim seems to be stronger than this. Evidently Suppes thinks that causes actually are dependent on a conceptual framework. This is different from the other theories we have discussed, in which causality was independent of our ways of looking at the world.

Rosen's first criticism of Hesslow is to point out that the conceptual framework from which causal relations are determined may be an inadequate or partial framework. Rosen never explains what would make one framework adequate, and another inadequate. One wonders why all conceptual frameworks are not equally adequate. More needs to be said about what makes a conceptual framework adequate or inadequate.

Assuming that sense can be made of the distinction between adequate and inadequate conceptual frameworks, let us look at Rosen's application of that idea.

Rosen notes that pregnancy (M) may cause thrombosis (T), the taking of contraceptive pills lowers the probability of pregnancy, and that the taking of contraceptive pills lowers the probability of thrombosis. The above is captured in the following relations: $P(T/M) > P(T)$, $P(M/C) > P(M)$, and $P(T/C) > P(T)$. Rosen (1978, p. 606) continues the discussion:

Consequently, based on the available information represented by the above probability estimates, we would be hesitant, where a person suffers a thrombosis, to blame the person's taking of contraceptive pills. But it does not follow from these epistemic observations that a particular person's use of contraceptive pills lowers the probability that she may suffer a thrombosis, for, unknown to us, her neurophysiological constitution (N) may be such that the taking of the pills definitely contributes to a thrombosis. Formally,

$$P(T/C,N) > P(T)$$

represents our more complete and accurate causal picture. We wrongly believe that taking the pills always lowers a person's probability of thrombosis because we base our belief on an inadequate and superficial knowledge of the causal structures in this medical domain where unanticipated and unappreciated neurophysiological features are not given sufficient attention or adequate weighting.

Rosen's reply is a claim that the taking of contraceptives lowers the probability of thrombosis in an inadequate and partial conceptual framework. She thinks that if we look at an adequate conceptual framework the taking of contraceptives will raise the probability of thrombosis. Rosen does not say which of the three, if any, of the conceptual frameworks that she mentioned we are dealing with in this case. It seems to me that we are dealing with the framework of our general beliefs that expresses all of the information available to us; one wonders why we would not want to use a framework that expresses all of the information available to us. Rosen's claim is that unknown to us, there may be other factors present such that the taking of contraceptives does raise the probability of thrombosis in these cases. This reply is interesting, because it seems to presuppose that the conceptual framework of all the information available to us is inadequate and partial. But this conceptual framework is one of the ones that Suppes and Rosen recognize as adequate. This conflict seems to suggest that any conceptual framework that allows a cause to lower the probability of its effect will be decreed inadequate and partial by Rosen. If this is true, then it should be recognized that a necessary condition

for a conceptual framework being adequate is that relative to it all causes raise the probability of their effects. This a priori requirement upon adequate conceptual frameworks seems unjustified to me, and on a par with the a priori requirement that all events have a sufficient cause. Whether or not causes always raise the probability of their effects, like the problem of determinism, appears to be an empirical question, and not one that can be settled a priori. We must realize that our standard of what constitutes an adequate conceptual frame is not dependent upon whether all causes raise the probability of their effects relative to it.

We thus see that Rosen's rejection of the conceptual framework in which contraceptives cause thrombosis must be based upon other reasons than the fact that in it contraceptives both can cause thrombosis and lower the probability of thrombosis. The reason Rosen gives is that there *may* be other factors present that complete the causal picture in such a way that contraceptives do raise the probability of thrombosis. Before we look at the justification for this claim, let us consider a minor point. Rosen says that there may be a neurophysiological condition N such that $P(T/C,N) > P(T)$. If this were true, then the *prima facie* cause, according to Suppes, would be the conjunction of N and C; the cause would be the neurophysiological condition and the taking of contraceptives. If Rosen wants to claim that the taking of contraceptives is a *prima facie* cause she must claim that $P(T/C,N) > P(T/N)$, which is probably true in this case. This is a minor point, but deserves to be mentioned.

Rosen's main claim that there *may* be other factors present that complete the causal picture appears problematic. First of all, even though there may be other factors N that we do not know about such that conditional on them, contraceptives do raise the probability of thrombosis, there may also be other factors W that we do not know about such that conditional on both N and W contraceptives lower the probability of thrombosis. It may be the case that there is a W, such that $P(T/C,N,W) < P(T)$. This observation is actually a claim that the conceptual framework that Rosen is referring to may be inadequate and partial also. One might reply to Rosen that if we take an adequate conceptual framework, and conditionalize on W also, then we will see that contraceptives do lower the probability of thrombosis.

Rosen could reply to this criticism that sooner or later we will find the

ultimate conceptual framework, which includes a factor Z , such that C does raise the probability of T , relative to Z and all of the other factors: $P(T/C,N,W,\dots,Z) > P(T)$. We should realize that this commits Rosen to claiming that there is an ultimate conceptual framework, relative to which all causal statements should be made. This may be true, but it appears inconsistent with Suppes' and Rosen's claim that there are different conceptual frameworks which are adequate. For this reason, Rosen may claim that there is no ultimate conceptual framework which includes a factor Z that fulfills the above relations. The list of conceptual frameworks may be infinite, and it is possible that for each one relative to which contraceptives raise the probability of thrombosis, there is another one, inclusive of the former, relative to which contraceptives lower the probability of thrombosis. This possibility shows us that Rosen must claim that a point must be reached in the list of conceptual frameworks at which all more inclusive conceptual frameworks have the property that contraceptives raise the probability of thrombosis.

The preceding discussion shows that Rosen's defense rests upon the claim that there *might* be further conditions that we do not know about that provide a more complete conceptual scheme. Two basic problems arise from this. The first is that there may be no conceptual framework that is more adequate than the original one and relative to which contraceptives raise the probability of thrombosis. Perhaps the more adequate conceptual framework contains factors that do not change the probability relation, or perhaps there is no more adequate conceptual framework. Just as its possible that there may be a more adequate conceptual framework that fits Rosen's needs, its possible that there may not be a more adequate conceptual framework that fits her needs. Secondly, problems arise when we consider that there may be other more adequate conceptual frameworks than the one that Rosen is postulating and which contain information relative to which contraceptives lower the probability of thrombosis. These problems with Rosen's defense show that it is weak and rests upon a firm belief that causes do raise the probability of their effects.

6.2.2. Rosen and Improbable Consequences

Rosen (1978, p. 607) also provides an alternative means of answering Hesslow's example:

That the taking of contraceptive pills might cause a woman to suffer a thrombosis is curious or unlikely because it is not, from an intuitive standpoint, what we would expect to happen. In fact, where the woman knows that pregnancy increases the likelihood of thrombosis, she might take the pill for the very reason that she wishes to avoid that consequence. At the time of taking the pills, then, she would not admit that this is a likely cause of what she is trying to avoid. It seems to me that this feature of Hesslow's example needs to be stressed. We are faced with an unanticipated and unlikely outcome. The problem is how to accommodate this feature in a causal calculus. Far from threatening Suppes' general approach, it seems to me that this general problem can only be explicitly accommodated by a probability approach. By using the technique of conditionalization to relativize probabilities to background information, a standard probability framework is the only one that can handle in an explicit way the complicating but familiar problem of events with improbable consequences.

In the previous defense of Suppes, Rosen admitted that the contraceptives were a cause of thrombosis, but claimed that contraceptives did not really lower the probability of thrombosis. In this defense of Suppes, Rosen admits that contraceptives lower the probability of thrombosis, but she denies that contraceptives are a cause of thrombosis. In this section, Rosen seems to think that the conceptual framework relative to which contraceptives lower the probability of thrombosis is a legitimate or adequate conceptual framework. This seems inconsistent with the previous claims that Rosen made. I agree that the woman taking the contraceptives would think it unlikely that the taking of contraceptives is a cause of thrombosis in this situation, but if she did develop a thrombosis, it seems as if she would say that the improbable happened, and the taking of contraceptives caused it, if she knew about the existing causal chain. Suppose that doctors could discover a direct causal chain connecting her taking of contraceptives to the thrombosis; in this case she would have to admit that the contraceptives were a cause of thrombosis, even though she had taken them to avoid thrombosis.

Considerations such as this one make the appeal to conceptual frameworks appear to add a subjective element into causation. Suppose that the woman above did not know about the direct causal chain linking the contraceptives with the

thrombosis; for her, the taking of the contraceptives is a negative cause of thrombosis. For her friend, knowing about the causal link, the taking of the contraceptives is a cause of thrombosis. Whether the taking of the contraceptives is a cause of thrombosis or not seems to depend on what one knows. This is a very subjective account of causation, which many may not be inclined to accept.

I find Rosen's defense of Suppes quite puzzling. It appears as if her two main lines of defense are inconsistent with one another. We must decide if we are going to require that we use only adequate conceptual frames, and if so, what is going to count as an adequate conceptual frame.

6.2.3. Rosen's Original Example

Rosen (1978, pp. 607-8) offers another example in an attempt to clarify her position. This example was one of the original examples which raised the problem of positive relevance.

Consider a mediocre golfer, Jones, who on her approach shot hits a tree-limb with the spectacular result that her ball is deflected directly into the hole for a birdie. In advance of play, we would give Jones a low probability of making a birdie on this hole. And we would estimate the probability of making a birdie as still lower given the information that Jones will hit a branch. What is even more curious is that when we see the event happen, we know immediately that had Jones not hit that limb in just the way that she did, the ball would not have gone into the cup. Similarly, to refer back to our earlier example, the woman's neurophysiological condition along with her taking contraceptives is just what gives her a thrombosis.

In this example we have a situation in which the improbable happens. Hitting the branch lowers the probability of making a birdie, and so it is not even a *prima facie* cause for Rosen and Suppes. The problem is that hitting the branch is in the causal chain connecting the approach shot and the birdie. Suppes and Rosen caution us against being content with a narrow view. Like the thrombosis example, Rosen attempts to handle this example by specifying the situation more closely in a manner that makes the hitting of the tree limb raise the probability of a birdie. What Rosen has done is specify a new narrower reference class in which hitting the tree limb is a *prima facie* cause of a birdie. We might characterize Rosen's method as one which finds a narrower reference class in which the event in question is a

prima facie cause. Salmon (1980, p. 64) refers to Rosen's method as the *method of more detailed specification of events*.

This method seems reasonable in the golf example, because we generally believe that if we specified the events precisely enough, the hitting of the branch would indeed raise the probability of a birdie. In fact, we know how we would specify the events more precisely: we would give the momentum and angle of deflection with the branch, and other variables in such a way that each state of the chain is a cause of a later state of the chain. This is different from the thrombosis case. In the thrombosis case, we did not really know how to specify the events more closely, and Rosen was forced to simply *assume* that if we did specify them closely enough that there would be neurophysiological factors that would make the contraceptives a prima facie cause of thrombosis. It is not obvious that specifying a narrower reference class by specifying the events precisely will always solve the problem of causes that do not raise the probability of their effects. For some examples, Rosen must simply have faith that there are relevant factors that are unknown.

6.2.4. Objectively Homogeneous Reference Classes and the Precise Specification of Events

Suppes proposes his account of probabilistic causation to account for the way in which we ordinarily use the idea of causation. But in order for it to do so, the reference classes from which we take the probability of certain events must not be required to be homogeneous reference classes. But if we don't have an objectively homogeneous reference class, it is very possible for an event to be a negative prima facie cause with respect to the larger reference class, and be a prima facie cause with respect to the narrower reference class. Rosen has used this tack to escape from saying that causes can lower the probability of their effects; when the event is placed in an objectively homogeneous reference class, or even a narrower reference class, it becomes a prima facie cause and raises the probability of its effect. But by using this method Rosen practically admits that in the ordinary use of the term "cause," a cause can lower the probability of its effect. We normally do not specify events precisely and require that homogeneous reference classes determine the probability of events, but rather we specify events loosely and assign probability

values on the basis of reference classes made up of events we consider similar in some respect. But then it does not seem so unusual that a cause lowers the probability of its effect, because we are using very loosely constructed reference classes to assign probability values. Rosen may be correct in saying that, if we specify events precisely and use objectively homogeneous reference classes, a cause always raises the probability of its effect. But if we are not specifying events precisely, a cause does not always raise the probability of its effect. If Suppes wishes to hold onto the idea that a cause always raises the probability of its effect, he must restrict his theory to specifying events precisely, determining the probability of events with objectively homogeneous reference classes, and forego attempting to capture the ordinary usage of language.

The question then remains whether, if we specify events precisely, a cause always raises the probability of its effect. Consider the following theorem:

Theorem 2: If B is a *prima facie* cause of A, B is a *prima facie* negative cause of $\neg A$, i.e.:

$$P(A/B) > P(A) \equiv P(\neg A/B) < P(\neg A).$$

Suppes is unclear about whether he is dealing with events or types of events, and thus he does not endorse either view. If we are dealing with events, if A occurs, then $\neg A$ does not exist, which raises problems for theorems such as this one. But for this discussion, I will go along with Suppes and also not endorse either position. This theorem is important in situations in which a *prima facie* cause occurs but the effect does not occur. Suppose that we have a true statistical law and that B is a *prima facie* cause of A. Since this is a statistical law, there will be times in which B occurs and A does not occur, or times in which B occurs and $\neg A$ occurs. But since B is a *prima facie* cause of A, it is a *prima facie* negative cause of $\neg A$. Thus we have B and $\neg A$ occurring, where B lowers the probability of $\neg A$. The important question in these instances is whether B is a cause of $\neg A$, or whether $\neg A$ has some other cause or no cause at all. Someone who firmly believed that causes raise the probability of their effects might claim that in situations like this we do not have a causal chain. It might be claimed that there is no causal process connecting the negative cause and the effect, or that possibly they are connected by a process, but these are not causal processes. This would be a way of claiming that all causes are positively relevant to their effects.

6.3. Indeterminism and the Improbable

Let us suppose for now that mark transmission is a sufficient condition for there to be a causal process. It certainly seems reasonable that if a mark is able to be transmitted, that causal influence is also being transmitted. This does not mean that every event in the process is caused by the preceding events in the process; it only means that some causal influence is transmitted. Now consider a causal process or chain which has some indeterministic links. It seems reasonable to believe that if the improbable occurs at some point in the process or chain, that the mark will still be transmitted. If this occurred, I think that we would have to say that there is a member of the causal process that lowers the probability of the next member in the process. I would say that this is a causal process because the mark is still present. For example, suppose that when two billiard balls collide, the outcome is indeterministic in the sense that very rarely they will travel in a direction not predicted by classical mechanics. Suppose we mark one of the balls, they collide, the unusual happens, and they travel in the direction that is predicted with low probability. In this situation, I think that we would have a causal process, even though some improbable events occurred. The mark is still transmitted, even though the unlikely event happened. If it were a true statistical law that the billiard balls behave in such a manner, there would have to be times in which the balls did behave in the way that is unlikely. When they behave in that manner, there seems no reason to think that the mark will disappear, or that this is not a causal interaction. We can describe the situation in such a way that it supports counterfactuals, which is also evidence that we have a causal interaction. This shows that if there are indeterminate events, there will probably be causal processes in which some events lower the probability of the later events.

This result shows that Reichenbach's hope of defining the causal net in terms of probability relations is impossible. Since there will be causal processes in which some events are negatively relevant to later events, we cannot define causal betweenness in terms of increasing probability. The only way to avoid this result would be to claim that whenever the improbable happens in an indeterminate situation, that a mark cannot be transmitted. This seems to me to be a very implausible assumption. Thus we must conclude that if indeterminism is true, there will be instances of causal processes in which each member does not have to be positively relevant to the next member.

6.4. I. J. Good and Positive Relevance

6.4.1. Good's Method of Interpolated Causal Links

A different approach to dealing with the problem of causes that are negatively relevant to their effects is provided by Good. It is a theorem of Good's theory that if any event in a chain is not positively relevant to the next link, then the strength of the whole chain is zero. Good (1961, p. 311) notes that this is equivalent to saying that there is no causal chain in this case. Thus the previous counterexamples are also counterexamples to Good's theory. Good's solution to the problem is found in a short note (1962, p. 88), which says that "it is worth noting that a 'cut' chain can often be uncut by filling it in more detail." The idea Good is expressing is that sometimes the spatio-temporal separation between events in a chain is too large, and may result in the appearance of a cut chain. But when the intermediate events are filled in, we see that the chain was not in fact cut, and that each event in the chain is positively relevant to the next event in the chain. This result led Good to require that the strength of a causal chain be the limit, as the size of the events approach zero, of the strengths of the chains. Although axiom 9 does not require that as the size of the events becomes zero, that more intermediate events fill in the space that the previous larger event occupied, it does seem that Good intended this to be the case. This way of dealing with the problem has been called *the method of interpolated causal links* by Salmon.

The method of interpolated causal links, like the method of more complete specification of events, seems to rest upon the belief that the omission of some relevant facts about the situation is responsible for the appearance that some causes do not raise the probability of their effects. Due to a faulty description that does not contain all of the relevant evidence, it appears as if some causes lower the probability of their effects. These methods seem to have faith that if we could give a proper description of the situation, that we would find that all causes raise the probability of their effects. This is analogous to the faith of the determinist who postulates some hidden variables when confronted with situations that appear to be indeterministic. However, there are two problems that seem to cast doubt upon the success of these methods.

6.4.2. Salmon's Counterexample

The first problem is a counterexample provided by Salmon. This example is a problem for the method of interpolated causal links because there does not seem to be any way to fill in the intermediate events in the chain. Salmon's (1980, p. 65) example is as follows:

We have an atom in an excited state which we shall refer to as the 4th energy level. It may decay to the ground state (zeroeth level) in several different ways, all of which involve intermediate occupation of the 1st energy level. Let $P(m \rightarrow n)$ stand for the probability that an atom in the mth level will drop directly to the nth level. Suppose we have the following probability values:

$$\begin{array}{ll} P(4 \rightarrow 3) = 3/4 & P(3 \rightarrow 1) = 3/4 \\ P(4 \rightarrow 2) = 1/4 & P(2 \rightarrow 1) = 1/4 \end{array}$$

It follows that the probability that the atom will occupy the 1st energy level in the process of decaying to the ground state is $10/16$; if, however, it occupies the 2nd level on its way down, then the probability of its occupying the 1st level is $1/4$. Therefore, occupying the 2nd level is negatively relevant to occupation of the 1st level. Nevertheless, if the atom goes from the 4th to the 2nd to the 1st level, that sequence constitutes a causal chain, in spite of the negative statistical relevance of the intermediate stage.

For the purposes of this fictitious example Salmon assumes that it is impossible for the atom to drop from the 3rd level to the 2nd level. This example is problematic, because there is no way, in principle, of filling in the missing links in the chain. There are no intermediate levels for the atom to occupy, and thus there are no intermediate links in the chain to fill in. Salmon also points out that this is a counterexample to the method of more detailed specification of events. There are no relevant facts that are not specified in the example.

6.4.3. Good's Reply to Salmon

Good replies to Salmon's example by questioning whether we actually have a causal chain in this situation. Clearly, if one denies that this is a causal chain there is no problem of causes that lower the probability of their effects. It seems evident though, that some processes of the sort mentioned by this example are causal processes. Marks can be transmitted, which is strong, if not conclusive,

evidence that we have a causal process. Whether or not all causal processes are causal chains may be questionable, but we would grant that this is a causal process. We will later discuss whether all events in a causal process are causes, or members of a causal chain.

Good does attempt to give another answer to the problem and he claims that Salmon has misdescribed the situation. Good's statement of this objection is very vague and it is difficult to see what the objection really is. Salmon reports that Good is claiming that the effect is more probable in the presence of a probabilistic cause than it is in the absence of any cause at all. Salmon's interpretation of Good on this matter is better than any that I have arrived at, and I think it should be accepted until evidence is produced to the contrary. If this is the correct interpretation of Good's objection, then it does seem to be inadequate. As Salmon (forthcoming, p. 236) notes, "this observation amounts to the assertion that an event is more likely to occur in circumstances in which it is possible than it is in circumstances (the absence of any probabilistic cause) in which it is impossible."

6.4.4. Humphreys' Theorem

The second problem that arises for the method of interpolated causal links is due to a theorem proved by Paul Humphreys. If we let an arrow stand for a probabilistic causal connection and Π be a product operator, Humphreys' (1980, p. 308) theorem is:

Theorem 3: Let $C_0 \rightarrow C_1 \rightarrow \dots \rightarrow C_n$ be a two state chain with the markov property, where $0 < P(C_i) < 1$, and $n \geq 1$. Then

$$P(C_n / C_0) - P(C_n) = \prod_{i=0}^{n-1} [P(C_{i+1} / C_i) - P(C_{i+1})] \prod_{i=1}^{n-1} P^{-1}(\neg C_i).$$

This theorem has some important consequences for two state Markov chains, which Good's theory applies to. This theorem tells us that whether or not an event is positively or negatively relevant to a later event depends upon the product of the positive or negative relevance of all of the individual links in the chain.. Thus if we have a chain in which the first event is negatively relevant to the later event, then an odd number of the links are negative links. As Salmon has noted, this is directly applicable to the method of interpolated causal links. If in the original chain the first event is negatively relevant to the last event, and if any more

detailed chain will also be a two state Markov chain, then any more detailed chain will also have at least one negative link. Humphreys' theorem shows that if every link in a chain is positive, the whole chain is a positive chain. Since Good's theory is applicable to two state Markov chains, we must conclude that his method of interpolating causal links is unsuccessful.

Humphrey's theorem assumes that the causal chain consists of n events, where n is a finite number. If we assume that the number of events in the chain is infinite, or is a continuum, then Humphreys' theorem is not applicable. Good and Rosen both assume that causal chains are finite. However, there does seem to be some reason for believing that some causal chains have a continuum of events. If we consider an arrow in flight, or a billiard ball in motion, it seems plausible to consider each point they occupy at a time to be an event. This seems compatible with Salmon's At-At theory of causality. The reduction of events in size to point events will enable the method of interpolated causal links to avoid the problem with Humphreys' theorem.

6.5. Salmon On Positive Relevance

6.5.1. The Method of Successive Reconditionalization

One further method of dealing with the problem of negative statistical relevance has been developed by Salmon. Salmon's method, which he calls *the method of successive reconditionalization*, contains parts of both the method of more precise specification of events and the method of interpolated causal links. The intuition behind the the method of successive reconditionalization is that alternative events that might have happened, but didn't, should be irrelevant to the determination of whether an event is a cause of another event. This method solves the problem by noticing that in a causal chain, each event is usually positively relevant to the next event, if we conditionalize on all of the preceding members of the chain; alternative events in chains that did not occur are not to be included in the reference classes of the probabilities. As a causal chain becomes actual, the reference class by which we determine if a cause raises the probability of its effect becomes narrower.

This method handles Rosen's golf example in a satisfactory manner. Swinging at the ball raises the probability of the ball being hit by the club; once the club has been swung, hitting the ball raises the probability of the ball traveling towards the branch; once the ball has been hit, the ball traveling toward the branch raises the probability of the ball hitting the branch; once the ball is traveling toward the branch, the ball hitting the branch raises the probability of being deflected toward the hole; and once the ball hits the branch, being deflected toward the hole raises the probability of it falling in the hole. This example illustrates how the method of successive reconditionalization narrows the reference class and changes what is conditionalized on as various events happen, and how each event in the chain is positively relevant to the next event in the chain. Salmon has shown in detail how this method also handles the tetrahedron-card game case. Although I have presented an informal account of this method, the reader can find a more detailed treatment of it in Salmon (1980).

6.5.2. Successive Reconditionalization and Indirect Causes

One nice benefit of the method of successive reconditionalization is that certain problems with indirect causes are solved. In Suppes' definition of an indirect cause, an indirect cause was required to be a *prima facie* cause, and thus raise the probability of its effect. The problem was that many events in a causal chain that we would consider indirect causes do not raise the probability of the final event. The method of successive reconditionalization does not require that indirect causes be positively relevant to the final effect; all that is required is that each cause in the chain be positively relevant to the next member of the chain. If one desired, a definition of indirect causes could be given which requires that an indirect cause be positively relevant to its immediate effect, instead of to the final effect. As Salmon (forthcoming, p. 233) notes, this also affects Reichenbach's idea of causal betweenness. No longer would it be necessary to require that all of the events in the causal chain be positively relevant to the end event. Because of this, one might be able to construct a better definition of causal betweenness by using the definition of causal relevance. We might say something like B is causally between A and C if A is causally relevant to B, and B is causally relevant to C. This would not have been satisfactory to Reichenbach, because the definition of

causal relevance assumes that a temporal direction has already been given. But it is certainly to the advantage of the method of successive reconditionalization that it solves this problem of indirect causes.

6.5.3. Problems With the Method of Successive Reconditionalization

Although the method of successive reconditionalization seems to work with normal causal chains, it will not work on all causal chains. It will not work with the example of the atom that we discussed earlier, because there is no way to specify the situation more precisely and narrow the reference class in the appropriate manner. In Rosen's golf example we added some links in the chain and conditionalized on them; in the example of the atom decaying, there are no more links in the chain that we can add. One feature of the atom is that the chain connecting occupation of the 4th energy level and the 2nd energy level has no intermediate links. This is quite different from the continuous chains connecting the events in the golf example, and has the result that the method of successive reconditionalization cannot handle this example.

The method of successive reconditionalization will not be able to handle any example in which the events are discrete in the sense that we cannot find intermediate events. This method relies heavily upon being able to conditionalize on intermediate members of the causal chain, but if there are no intermediate members of the causal chain, the method will fail. Suppose that A is a probabilistic predecessor of C, and of $\neg C$; by this I mean that sometimes A will cause C and other times A will cause $\neg C$. Let us also suppose that either C or $\neg C$ can cause E, but that C lowers the probability of E, and A lowers the probability of C. If we have a causal chain consisting of A, C, and E, the method of successive reconditionalization tells us to look for intermediate events to conditionalize on. But if there are no events in those locations, then this method will fail. Most causal chains that we notice are continuous, and it appears that this method will work for them. But in the quantum domain, where causal chains are not always continuous, this method will fail.

6.6. Causes, Causal Factors, and Positive Relevance

6.6.1. Causes and Negative Causes

Since we must conclude that efforts to preserve positive statistical relevance for causes have ultimately failed, it will profit us to return to the original problem and decide if positive statistical relevance is really necessary. Suppes introduced the idea of a negative cause, which is supposed to capture the idea of a cause that tends to prevent an event from occurring. He (1970, p. 43) gives the following definition of a negative cause:

Definition 4: The event B_i is a *prima facie* negative cause of A_i if and only if

1. $t' < t$,
2. $P(B_i) > 0$,
3. $P(A_i / B_i) < P(A_i)$.

One wonders why Suppes was so concerned about showing that each member of the causal chain in Rosen's golf example was positively relevant to the effect, if he also believed in the existence of negative causes. The only explanation that seems reasonable is that even though Suppes does believe in negative causes, he believes that they are never connected by causal processes to the effect. In other words, some causally relevant conditions need not be connected to the effect by means of causal processes. This is in contrast to Reichenbach's idea which was that causes simply were the events that are connected to other events by causal processes. Suppes might hold that all events that are connected to other events by causal processes are positive causes of those events, and hence positively relevant to them, but that other events which are not connected to the effect by a causal process can be negatively relevant to the effect.

6.6.2. Causes That Are Not Causal Factors

If we look closely we find that many events are causes of other events, even if there is no causal process connecting them. This will occur whenever a negative event is a cause of something. There will not be a causal connection between a negative event and the effect, but we often classify it as a cause. Consider the following example. Suppose the local fire department is on strike, and is refusing to respond to requests to put out fires. If a house catches on fire and burns to the ground, it is natural to say a cause of the house burning to the ground is that the fire department is on strike. There is no causal process connecting the fire department being on strike with the house burning down, and yet we still consider it a cause. If the fire department had not been on strike, the house would not have burned to the ground. Other examples similar to this abound. A friend's failure to keep an appointment is a cause of my being late for a show, the ambulance's slowness in responding to a call caused a person to die, etc. Events such as these support counterfactuals, and are classified as causes. These negative events are also negative causes. Thus by drawing the distinction between causes, which do not need to be connected with their effects by processes, and causal factors, which simply are events connected to another event by a causal process, we are able to see the motivation behind Suppes' defense of positive relevance in Rosen's golf example.

Having seen that not all causes are causal factors, the next question is whether all causal factors are causes. Most philosophers, with the exception of Humphreys, have thought that all causal factors are causes. If we simply define a causal factor of an event to be an event that is connected to another event by a causal process, we must claim that not all causal factors are causes. The reason for this is because it is possible for events to be connected by a causal process even though there is no causal influence being transmitted through any interaction with the process. Causal processes can intersect without interacting. This will be discussed in more detail in a later chapter. But we must now discuss Humphreys' claim that not all causal factors are causes.

6.7. Humphreys' Theory of Causality

6.7.1. Humphreys' Example

One possible way to claim that causes are always positively relevant to their effects is to claim that not every event in a causal chain is a cause. One might claim that causal chains are made up of both negative and positive causes, but only positive causes are properly called causes. Humphreys takes a position similar to this. Humphreys (1980, pp. 407-8) presents the following example:

Two climbers at the top of a mountain have had their normal route back cut off by an avalanche. The only other way down is via an easy rock face with one very dangerous overhang. Once over that, the route is clear to the bottom, but inexpert climbers are likely to fall at the overhang. A fall, of course, increases the probability of injury at that point, and if anyone is injured at the overhang, it is highly unlikely that he could complete the journey. Of the two climbers, Axel is far superior to his companion. So it is decided that Axel will go for help.

In this example, Axel being chosen to go for help is positively relevant to getting help. However, sometimes the unexpected will happen, Axel will fall at the overhang, and the unexpected happens again when he fails to hurt himself. In this case we have a causal chain connecting Axel's being chosen, Axel's falling, Axel's not getting hurt, and Axel's getting help. But even though there is a causal chain connecting these events, Humphreys does not think that Axel's being chosen is a cause, or even an indirect cause, of help coming. Humphreys (1980, p. 408) says:

On this occasion, should we say that A (Axel being chosen) caused D (climber fetches help)? I think not. Axel's skill was *prima facie* relevant at the overhang, yet it played no role in the final outcome as events turned out. It was certainly not his skill which led to his slipping nor to his escaping injury; in fact it was the combination of two relatively improbable events which brought about the final result. So in this sequence of events, the fact that Axel was chosen had nothing to do with help being fetched, despite the fact that in general, it would have. The conclusion we must reach is that we cannot decide, without examining the whole process between the two events, whether non-spurious *prima facie* causes are genuine causes.

Humphreys thinks that Axel's being chosen was not a cause of help being fetched, because there are some negative links in the chain. The negative links prevent Axel's being chosen from being a cause, even though there is a causal process present.

6.7.2. Critical Discussion of Humphreys' Theory

In this example, I think that it is plausible to suppose that if we used the method of successive reconditionalization at each stage in the process, then we would have a causal chain in which each event is positively relevant to the next event. Since I don't know the details of Axel's performance at the overhang, I will not attempt to fill in any details and complete the causal picture. But if we assume that it is possible to reconditionalize and make each event in the chain positively relevant to the next event, an interesting problem arises. Given that Humphreys does not think that Axel's being chosen to go is a cause of help arriving in the situation in which he described, it seems unlikely to me that he would think that Axel's being chosen is a cause of help arriving if we are able to redescribe the situation in such a way that each event in the causal chain is positively relevant to the next. I am not sure about this, and Humphreys (1980, p. 311) does warn us against not giving a complete description of the situation.

I think that some insight can be gained into this situation if we contrast it with an analogous situation that we construct from Rosen's golf example. Suppose that we have two golfers, Arnold and Rod, who are partners in a golf contest. One of them has to hit the ball, and if he makes a birdie, they win the prize. Since Arnold is a rather good golfer, and Rod is a rather poor golfer, they decide to let Arnold try for the birdie. Unfortunately, Arnold slices the ball, which then hits the tree, and finally lands in the hole for a birdie. In this situation, it was improbable that Arnold would slice the ball, and that the ball would go in the hole after hitting the tree. Thus there are negative links in the chain connecting Arnold's being chosen to hit the ball, and the ball going into the hole. This example is analogous to Humphreys' mountain climbing example, with the difference being that we know how to fill in the details in this case to make each event positively relevant to the next event in the chain. As we have seen previously, we can use the method of successive reconditionalization to see that each event in the chain is positively relevant to the next event in the chain. Although I supposed this could be done in Humphreys' example, I did not attempt to do it.

Humphreys claimed that Axel's being chosen was not a cause of help arriving because his skill did not affect the outcome. The final outcome was the result of two rather improbable events instead of Axel's skill, and thus Axel's being chosen is

not a cause of help arriving. Similarly, we might claim that Arnold's being chosen to hit the ball was not a cause of the ball going in the hole, because the ball going in the hole was not a result of Arnold's skill, but rather it was due to some rather improbable events. The reasoning between these two cases is parallel. Thus it seems as if Axel's being chosen is not a cause of help arriving, then Arnold's being chosen is not a cause of the birdie.

Contrary to Humphreys' reasoning, it seems to me that Arnold's being chosen is a cause of the birdie. Because it is possible to show, through the method of successive reconditionalization, that Arnold's being chosen did have something to do with the birdie, it seems as if it is a cause of the birdie. Humphreys claimed that Axel's being chosen had nothing to do with help arriving, and thus he is hesitant to claim that it is a cause of help arriving. It seems to me that this claim is false, if the method of successive reconditionalization will work on this example. I see no reason to think that this method will not work on this example, because it seems to work on all cases of macroscopic causation. Thus we must conclude that if Humphreys really does have the intuition that Axel's being chosen is not a cause of help arriving, it cannot be for the reasons he gives.

Perhaps the reason that some may be tempted to say that Axel's being chosen is not a cause of help arriving is because Axel was very lucky. Help arrived because of luck, and not because of Axel's skill; we might be hesitant to give Axel the credit for help arriving. Similarly, we would say that Arnold got lucky; Arnold's winning the prize was due to luck and not due to skill. If I were Arnold's competitor, I would probably resent his winning. It seems as if in dealing out credit or blame in matters such as this, the actual existence of causal chains in which each event is positively relevant to the next event is not as important as the ability to foresee the actual causal chain, and intend that it occur. It also seems as if the statistical relevance of the cause to the final event is relevant in matters of guilt and blame. If we were to consider examples similar in structure to Humphreys' example, but did not have any human element of intending present, I think that we would consider the first event a cause of the final event, even though some improbable things happened. As long as each event is positively relevant to its successor, it appears as if the first event can be a cause of the final event. Outside of the quantum domain, it is not clear whether there are ever any causal chains in which each event is not positively relevant to the next event.

Aside from the previous problems, there are simple counterexamples to show that Humphreys' definition of causation is too broad. Humphreys says that C is a cause of E if there is a causal chain connecting them in which each event is positively relevant to the next event. He also requires that none of the connections be spurious connections. But on this account, it is possible that someone's birth is a cause of their death. It seems reasonable that at least some people, if not all, go through life in such a way that each event in their life is positively relevant to the next event in their life. But then it will be the case that being born is a cause of dying, which seems strange (van Fraassen 1980). This example shows that Humphreys' analysis of causation is too broad.

It is also easy to see that Humphreys' definition of a cause is too narrow. As we saw earlier, many events are causes of other events which are not connected to those events by processes, as Humphreys' analysis requires. Thus many genuine cases of causation will be excluded by Humphreys' analysis, which makes it too narrow.

6.8. Causes and Causal Factors

The previous problems arise for Humphreys because he is interested in picking out causes, instead of causal factors. Humphreys did notice that if a causal factor is just any event that is in the causal chain leading to the event, that not all causal factors have to be causes. But it is also true that not all causes are causal factors. It is clear that Reichenbach was not concerned with causes; he was attempting to use probability relations to construct the causal net of the world, which is the connections between all of the causal factors. It also seems as if Good was concerned with causal factors and not with causes. It is not clear what Suppes was concerned with, although it appears as if he was interested in causes and not actual causal connections among causal factors.

The problems of picking out causes and picking out causal factors are separate problems. It seems clear that positive statistical relevance is not required for causal factors. Salmon's example shows that there can be negatively relevant events in a causal chain, even when we use the method of successive reconditionalization. Thus it seems as if it will be impossible to pick out causal

factors using positive statistical relevance. It is not clear whether positive relevance is required for causes. I am inclined to think that some sort of positive relevance is required for causes, because it seems to be part of the meaning of a cause that it help bring about its effect. I think that most philosophers who have written about probabilistic causality have not seriously considered the difference between causal factors and causes. This resulted in the requirement of positive relevance for causal factors, which seems to be a false requirement. When discussing probabilistic causation, we must first become clear on what we are attempting to do: are we dealing with causal factors or with causes? The answer to this question will be relevant to whether we require positive statistical relevance.

CHAPTER 7

SIMPSON'S PARADOX

7.1. Probabilistic Causality and Simpson's Paradox

Nancy Cartwright (1979) has written a recent article in which she discusses the basic idea of probabilistic causality, namely, that a cause raises the probability of its effect. We have seen that all theories of probabilistic causality that we have discussed assume that a cause raises the probability of its effect. Cartwright questions this assumption, by noting that it is often false that a cause raises the probability of its effect. We saw an example of this that was given by Hesslow. Cartwright asks us to consider the following example. Let us suppose that smoking causes heart disease, and that exercise prevents heart disease. Cartwright uses an arrow to symbolize causal laws; thus we symbolize smoking causes heart disease by $S \rightarrow H$, and we symbolize exercise prevents heart disease by $X \rightarrow -H$. Since we believe that smoking causes heart disease, we would expect smoking to raise the probability of heart disease, $P(H/S) > P(H)$. However, this may not be true if smoking and exercise are highly correlated, and if exercise is more effective at preventing heart disease than smoking is at causing it. For example, suppose that people who like to smoke also like to exercise, and that the result is that they have a lesser chance of having heart disease than a normal person. In this situation we find that $P(H/S) < P(H)$, because most of the smokers are also people who exercise, and exercise is better at preventing heart disease than exercise is at causing it. In this situation we find that although smoking causes heart disease, smoking actually lowers the probability of heart disease.

A situation such as this will arise anytime the cause is sufficiently correlated with a third preventative factor of sufficient strength. This is known as Simpson's paradox. Simpson's paradox is that any correlation in a population can be reversed in the subpopulations by finding a third variable that is properly correlated with the other variables. Thus positive relevance can become negative relevance or independence, negative relevance can become positive relevance or independence, and

independence can become positive or negative relevance. Almost anything is possible. But if this is true, it seems hopeless to require that causes raise the probability of their effects. By looking at the proper reference class we can reverse relevance relations, so it appears arbitrary to claim that causes raise the probability of their effects.

7.2. Cartwright's Treatment of Simpson's Paradox

7.2.1. Informal Statement of Cartwright's Solution

In response to this disturbing fact, Cartwright notices that all of the counterexamples to the claim that causes are positively relevant to their effects depend upon a correlation between the cause and some other causally relevant factor in such a way that the positive relevance is changed. The natural answer to these counterexamples, which Cartwright sees, is to claim that causes increase the probability of their effects when these correlations are absent. The problem is to try to characterize this more explicitly.

In the example we discussed concerning smoking, exercise, and heart disease, we saw that smoking lowered the probability of contracting heart disease. This was because most smokers were also exercisers. Exercising is the causally relevant factor which reverses the positive relevance of smoking to heart disease. However, Cartwright notices that in situations in which there are none of these correlations, smoking always raises the probability of heart disease. If we had a population in which everyone exercised, and thus smoking was not correlated with exercise, then we would find that smoking increased the probability of heart disease. Similarly, if we took a population in which nobody exercised, smoking would not be correlated with exercise. In this population we would also find that smoking increased the probability of heart disease. Cartwright believes that in populations which have no correlations between the cause and other causally relevant factors, a cause will always raise the probability of its effect. These populations which have no correlation between the cause and other causally relevant factors are called *causally homogeneous*. Cartwright (1979, p. 423) summarizes this principle by claiming that "C causes E if and only if C increases the probability of E in every situation which

is otherwise causally homogeneous with respect to E." This principle claims that if we take the reference class to be causally homogeneous, then a cause always raises the probability of its effect. This is an answer to the problem posed by Simpson's paradox. This principle also claims that if C increases the probability of E in every causally homogeneous reference class, then C is a cause of E; this gives a sufficient condition for a probabilistic cause.

7.2.2. A Formal Version of Cartwright's Solution

In formalizing the idea of the above principle, Cartwright appeals to the idea of a state description which Carnap introduced in his work in inductive logic. A state description is a total or complete description of a possible state of the universe describable by the language. A state description will be a conjunction of simple statements, and each simple statement will either claim that a certain individual has a simple property, or deny that a certain individual has a property. Assuming that we have a set of simple properties, a state description will assign to each individual some of those properties, and the complement of those properties that are not assigned to it. Thus a state description is a maximal description of the properties that each individual could consistently have.

Cartwright adapts this idea to talk about maximal sets of consistent causal factors. We need to find a way to characterize all of the situations which are causally homogeneous with respect to E. Cartwright begins by defining a complete set of causal factors for E to be the set of all C_i such that either C_i causes E, $C_i \rightarrow +E$, or C_i prevents E, $C_i \rightarrow -E$; this will be written as $C_i \rightarrow \pm E$. If we take a possible arrangement of factors of this set minus C, then we have a situation that is causally homogeneous in all causal factors except C. These possible arrangements are generated by conjoining either the members of the set or their complements. A state description K_j is defined as $K_j = \bigwedge \pm C_i$ over the set $\{C_i\}$ of all alternative causal factors, and where i ranges from 1 to n. This will give us 2^n state descriptions. Each state description will be causally homogeneous with respect to E; these are the populations in which a cause must always raise the probability of its effect.

Cartwright (1979, p. 423) then uses the above conception of state descriptions to give a general principle which defines probabilistic causality:

CC $C \rightarrow E$ if and only if $P(E/C \& K_j) > P(E/K_j)$ for all state descriptions K_j over the set $\{C_i\}$, where $\{C_i\}$ satisfies

$$C_i \in \{C_i\} \Rightarrow C_i \rightarrow \pm E \quad (7.1)$$

$$C_i \notin \{C_i\} \quad (7.2)$$

$$(D)(D \rightarrow \pm E \Rightarrow (D=C_i \vee D \in \{C_i\})) \quad (7.3)$$

$$C_i \in \{C_i\} \Rightarrow \neg(C \rightarrow C_i). \quad (7.4)$$

This principle is not an analysis of $C \rightarrow E$, because the causal relation appears on both sides of the equivalence in the above principle. Cartwright views principle CC as providing the connection between probability relations and causality.

When reading principle CC one is struck by the loose and non-standard notation that it uses. We could rewrite CC in standard notation as follows:

CC' $C \rightarrow E$ if and only if $P(E/C \& K_j) > P(E/K_j)$ for all state descriptions K_j over the set S such that

$$S = \{C_i | C_i \rightarrow \pm E \& C_i \neq C \& \neg(C \rightarrow C_i)\}.$$

Although this is more precise than principle CC, I will continue to discuss principle CC since it is the principle that Cartwright herself proposed.

In principle CC, condition (7.1) tells us that if anything is a member of $\{C_i\}$, then it is either a cause of $+E$ or of $-E$. There are no causally irrelevant events that are members of $\{C_i\}$. Condition (7.2) tells us that C is not a member of $\{C_i\}$. This condition is necessary, otherwise $P(E/C \& K_j) > P(E/K_j)$ would always fail. If the state description K_j contained C , then the two probabilities would be equal. If the state description contained $\neg C$, then $P(E/C \& K_j) = 0$, and the inequality would again fail. Thus condition (7.2) is necessary. Condition (7.3) tells us that if any event is a cause of $+E$ or $-E$, then it is a member of $\{C_i\}$, unless it is the event C . This condition ensures that $\{C_i\}$ will contain all causally relevant factors. Condition (7.4) is added to require that the state descriptions do not contain events in the causal chain between C and E ; if events in the causal chain between C and E were in the state descriptions, they might screen off C from E , in which case the above principle would be false.

It is important to realize that the causal law $C \rightarrow E$ is composed of general terms, and not specific events. C and E are not specific events that have occurred. Instead they are types of events. Thus we might say that smoking causes heart disease, even though not every case of smoking causes heart disease; this solves the

problems that we discussed in reference to Reichenbach and Suppes concerning whether the cause and effect actually occurred. Cartwright is not interested in specific causal chains.

7.2.3. Modifications of Principle CC

In looking closely at principle CC, one immediately notices that requirements (7.3) and (7.4) are inconsistent. This has the unfortunate result that there never are any true causal laws. Conditions (7.3) and (7.4) are inconsistent because (7.3) requires that any member of the causal chain between C and E be included in $\{C_i\}$, whereas (7.4) requires that they not be included in $\{C_i\}$. This difficulty can be remedied by changing (7.3) to:

$$(D)[D \rightarrow \pm E \Rightarrow (D=C \text{ or } C \rightarrow D \text{ or } D \in \{C_i\})]. \quad (7.5)$$

Condition (7.5) is not inconsistent with (7.4), because if an event is an effect of C and a cause of E, then (7.5) does not require that it be a member of $\{C_i\}$. This modification does not affect the rest of the requirements, and does not alter the meaning of what it is to be a causally homogeneous population.

Another problem for principle CC arises from the possibility that some of the state descriptions may be logically inconsistent. If a state description K_j is logically inconsistent, then $P(E/C \& K_j) = 0 = P(E/K_j)$. This would also imply that there could be no true causal laws. The reason that some of the state descriptions may be logically inconsistent is that condition (7.3) and (7.5) require that any event that causes or prevents E, with the exception of C or effects of C, must be a member of $\{C_i\}$. But suppose that C_i causes E, ($C_i \rightarrow +E$), and that $\neg C_i$ prevents E from occurring, ($\neg C_i \rightarrow -E$). In this case both C_i and $\neg C_i$ are required to be members of $\{C_i\}$. We constructed the state descriptions by letting $K_j = \Lambda \pm C_i$ over the set $\{C_i\}$, where i ranges from 1 to n. Thus if both C_i and $\neg C_i$ are members of $\{C_i\}$, then half of the state descriptions will be inconsistent. A quarter of the state descriptions will contain C_i and $\neg C_i$, and another quarter will contain $\neg C_i$ and $\neg\neg C_i$; all of these state descriptions are inconsistent. Thus we see that there can be no true causal laws, because for half of the K_j 's, $P(E/C \& K_j) = 0$.

This defect in CC can be avoided by not allowing both C_i and $\neg C_i$ to be members of $\{C_i\}$. Two modifications are needed to avoid this problem. The first modification is to replace (7.5) by condition (7.6):

$$(D)[D \rightarrow \pm E \Rightarrow (D=C_i \text{ or } C_j \rightarrow D \text{ or } D \in \{C_i\} \text{ or } (\neg D \in \{C_i\} \text{ and } \neg D \rightarrow \pm E))]. \quad (7.6)$$

This modification does not require D to be a member of $\{C_i\}$, if $\neg D$ is a member of $\{C_i\}$. However, replacing (7.5) by (7.6) does not totally solve the problem. We must also require that if C_i is a member of $\{C_i\}$, then $\neg C_i$ cannot be a member of $\{C_i\}$. This could be done in various ways, the easiest being to add requirement (7.7):

$$C_i \in \{C_i\} \Rightarrow \neg C_i \notin \{C_i\}. \quad (7.7)$$

Requirement (7.7) does not allow both C_i and $\neg C_i$ to be members of the basic set that we construct our state description from. These two modifications enable us to avoid the above criticism.

A problem with inconsistent state descriptions still remains, even with the above modifications. Suppose that C_i is a cause of E , that $\neg C_i$ is a cause of $\neg E$, and that C_i is a member of $\{C_i\}$. The above modifications ensure that $\neg C_i$ is not a member of $\{C_i\}$, but that is not enough to ensure that the state descriptions are logically consistent. Suppose that the extension of $\neg C_i$ is a proper subset of the extension of C_j , $\square(x)[C_jx \Rightarrow \neg C_ix]$. In this situation both C_i and C_j could be members of $\{C_i\}$, and yet the conjunction of C_i and C_j is still an inconsistent predicate: nothing could exemplify it. This too would have the result that there could be no true causal laws.

An easy way to avoid the above problem is to require that the causal factors be disjoint:

$$\text{for all } C_i, C_j \in \{C_i\}, \neg \square(x)(C_i x \Rightarrow C_j x). \quad (7.8)$$

This requirement requires that the extensions of all of the C_i that are members of $\{C_i\}$ and their complements must be disjoint. This prevents the extension of $\neg C_i$ from being a proper subset of the extension of C_j , which gave rise to the present objection.

A different problem arises if we consider what condition (7.4) was meant to rule out. We needed condition (7.4) to prevent a state description from fixing part of the causal chain between C and E . If C causes C_i and C_i causes E , then if our state description contained C_i , the addition of C might not make any difference to

the probability. To avoid this problem, Cartwright simply excluded those intermediate causes from $\{C_i\}$. However, this does not solve the problem. Suppose that C_i is in the causal chain between C and E. We know that C_i is not a member of $\{C_i\}$. Suppose further that $\neg C_i$ is a cause of $\neg E$; then by conditions (7.3), (7.5), and (7.6), we must have $\neg C_i \in \{C_i\}$. This creates two problems. The first problem arises when a state description contains $\neg C_i$. Since $\neg C_i$ is later than C, it seems reasonable to suppose that in some cases $\neg C_i$ will screen off C from E. Thus it will be false that $P(E/C \& K) > P(E/K)$, because the two probabilities will be equal. The second problem arises when a state description does not contain $\neg C_i$; in this case it will contain C_i . But then the inequality will fail, because C_i might screen off C from E. We thus see that if we allow the complement of a member of the causal chain between C and E to be a member of $\{C_i\}$, then we are forced to conclude that there are no true causal laws.

The obvious solution to this problem is to require that the complement of these effects of C not belong to $\{C_i\}$. Let us replace requirement (7.4) by (7.9):

$$C_i \in \{C_i\} \Rightarrow [\neg(C \rightarrow C_i) \& \neg(C \rightarrow \neg C_i)]. \quad (7.9)$$

This condition excludes both effects and complements of effects of C from being members of $\{C_i\}$. However it is not strong enough to prevent all effects that might screen off C from E. It is possible that if C causes C_i , that only a part of $\neg C_i$ would be able to screen off C from E. In order to avoid unnecessary formalism, I will simply assume that we have excluded all parts of $\neg C_i$ that might screen off C from E.

Cartwright discusses the fact that principle CC does not have any temporal requirements built into it, and she feels that it is desirable to leave any temporal considerations out of CC. Her reasoning for this rests not upon any hope of developing a causal theory of time, but rather upon the possibility of backwards causation. I also feel that it would be nice to be able to leave temporal considerations out of an analysis of causality, but that in the present discussion there are some compelling reasons for introducing them. If we do not modify the conditions presented so far, principle CC will always be false. Suppose that a predicate D is defined as $\square(x)[Dx \equiv (Cx \& Fx)]$, where F is some predicate consistent with and independent of C. In this situation, D will be a member of $\{C_i\}$. But then the inequality will fail for some K_j because $P(E/C \& K_j)$ will be

equal to $P(E/K_j)$. This example shows us that we can't let D be composed of C . Another similar problem is we define F as $\square(x)[Fx \equiv (D_1x \& D_2x)]$, where D_1 and D_2 satisfy the condition $\square(x)[Cx \Rightarrow (D_1x \& D_2x)]$. In this situation $C \notin \{C_i\}$, and yet in effect C will be a member of some of the state descriptions.

One way to avoid these problems is to introduce temporal constraints. If we require that every member of $\{C_i\}$ occur earlier than C , then these problems do not arise. This requirement also eliminates the need for conditions (7.9) and (7.2). C does not occur before itself, so condition (7.2) is not needed to exclude C from $\{C_i\}$. Any effect of C or the complement of any effect of C occurs later than C , and thus (7.9) is not needed to exclude them from $\{C_i\}$. If we do add this temporal restraint, we can drop conditions (7.9) and (7.2), but we will have to modify condition (7.6). Condition (7.6) must be modified to (7.10):

$$(D)[D \rightarrow \pm E \Rightarrow (D \in \{C_i\} \text{ or } (\neg D \in \{C_i\} \text{ and } \neg D \rightarrow \pm E) \text{ or } D \text{ does not occur earlier than } C)]. \quad (7.10)$$

This modification is necessary, because (7.6) was inconsistent with our new temporal requirement. Thus our modification of CC consists of principles (7.1), (7.10), (7.4), (7.7), and the other restraints mentioned.

Cartwright attempted to give a syntactic definition of the set S . Unfortunately, her definition of S was deficient and I introduced some semantic notions in order to save it. An easier way to characterize S would be to do it totally semantically, by requiring that all of the C_i 's that are members of S be logically consistent with one another; logically consistent predicates are predicates that can be co-exemplified. Thus the state descriptions could be defined as maximal logically consistent sets of causal factors. We would still have to add certain requirements to exclude C and its effects from the state descriptions, but it looks as if this might be easier than modifying CC syntactically.

7.2.4. Sufficient Causes and Necessary Causes

One problem with the conditions given so far that Cartwright discusses deals with alternative sufficient causes. Suppose that C_j is a sufficient cause of E . Then in some state descriptions $P(E/C \& K_j) = P(E/K_j)$, and principle CC fails. This example does not even depend upon there being just one alternative sufficient cause;

several other causes may become sufficient when conjoined together. C_j and C_k may both be probabilistic causes, but together a sufficient cause of E. In this case some state descriptions will contain both C_j and C_k , which will make the inequality fail, because $P(E/C\&K_j)$ will be equal to $P(E/K_j)$ when K_j contains both C_j and C_k . The easy way out of this problem is to deny that any sufficient causes exist. I find this solution unsatisfactory, for reasons that have been discussed earlier. Cartwright recognizes this problem, and gives the following solution. Cartwright (1979, p. 428) modifies the beginning of principle CC to read:

CC* $C \rightarrow E$ if and only if $(j)[P(E/C\&K_j) > P(E/K_j) \text{ or } P(E/K_j) = 1 = P(E/C\&K_j)]$ and $(\exists j)[P(E/K_j) \neq 1]$.

The second conjunct is needed to prevent everything from being the cause of a necessary fact. This modification of Cartwright's is sufficient to solve the problems that arise from alternative sufficient conditions.

Although Cartwright has successfully dealt with alternative sufficient causes, she has not dealt with alternative necessary causes. Suppose that C_j is a necessary cause of E; from this follows that C_j will be part of half of the state descriptions, and that $\neg C_j$ will be part of the other state descriptions. But for any state description K_j which contains $\neg C_j$, $P(E/C\&K_j) = 0 = P(E/K_j)$, and the inequality fails. We can easily modify CC to account for this problem:

CC** $C \rightarrow E$ if and only if $(j)[P(E/C\&K_j) > P(E/K_j) \text{ or } P(E/C\&K_j) = 0 = P(E/K_j) \text{ or } P(E/C\&K_j) = 1 = P(E/K_j)]$ and $(\exists j)[1 < P(E/K_j) < 0]$.

This modification will account for cases in which a state description contains a necessary cause of E. I believe that the previous modifications are all necessary in order to avoid serious difficulties that present themselves even before we discuss whether the intuition that CC tries to capture is correct.

7.2.5. Philosophical Problems With Cartwright's Solution

It appears to me that principle CC is too strong in that there are instances in which a cause does not raise the probability of its effect in every causally homogeneous situation. As an example, I would like to consider an example that Cartwright presents in support of her theory. Ingesting a poisonous acid is a cause of death, ingesting a poisonous alkali may also be a cause of death, but ingesting

both a poisonous alkali and acid may not be harmful at all. From this example, Cartwright (1979, p. 428) states three "causal truths:"

1. ingesting an acid without ingesting an alkali causes death;
2. ingesting an alkali without ingesting an acid causes death; and
3. ingesting both an alkali and an acid does not cause harm.

Cartwright claims that these three claims are consistent with CC. It may appear initially as if she is correct in that belief. However, a question arises as to whether CC is consistent with the original claim that ingesting a poisonous acid causes death, and that ingesting a poisonous alkali causes death.

In order for the ingestion of acid to be a cause of death, it must raise the probability of death in every causally homogeneous population. But some of these causally homogeneous state descriptions will contain the ingestion of alkali. In these state descriptions, the ingestion of acid will actually lower the probability of death instead of raising it as it should. Thus according to principle CC, the ingestion of poisonous acid is not a cause of death. Similar reasoning will show that according to principle CC, ingestion of poisonous alkali is not a cause of death; both of these results are unintuitive.

I suspect Cartwright would reply to this line of reasoning by claiming that the ingestion of acid poison is not a cause of death, but that ingestion of acid poison without ingestion of alkali is a cause of death. Similarly, ingestion of alkali is not a cause of death, but ingestion of alkali without ingestion of acid is a cause of death. I find this very unintuitive. In order to save the claim that a cause raise the probability of its effect in every causally homogeneous state description Cartwright has resorted to requiring that something similar to a total cause be specified. There will always be causal defeaters that will prevent a cause from raising the probability of its effect. I find it unacceptable to require that all of these causal defeaters be excluded in the cause. Cartwright herself introduced this example (1979, p. 428) by claiming that "ingesting an acid poison may cause death; so too may the ingestion of an alkali poison." It seems obvious that in this case the ingestion of acid does cause death, even though there are situations in which it does not raise the probability of death. This example shows Cartwright's basic intuition that a cause must raise the probability of its effect in every causally

homogeneous population to be mistaken. Although it may be true in most cases, it does not seem to be true in every case.

Let us consider another example in which principle CC requires that a total cause be given. Suppose that we have a laser connected by wires to a power supply. Turning on the power supply does seem to be a cause of the laser firing. This appears to be uncontroversial. However, let us see if turning on the power supply raises the probability of the laser firing in every causally homogeneous state description. One member of the set $\{C_i\}$ will be that there is a break in the wires connecting the laser to the power supply, or that before the power supply is turned on someone cut the wires. These events are causes that prevent the laser from firing, and thus will be members of $\{C_i\}$. But if the state description contains the information that the wires are broken, then turning on the power supply does not raise the probability of the laser firing; the power supply has no effect on the laser in this situation. It is important to notice that the probability of the laser firing in this situation is not zero; although it is extremely unlikely, it is possible that the laser could fire without power from the power supply. If the probability of the laser firing given that the wires were cut was zero, then principle CC would be able to handle this case. However, in this case the probability of the laser firing given that the wires are cut is not zero; thus principle CC is faulty. In this situation principle CC would tell us that the cause of the laser firing is turning on the power supply when nobody has cut the wires, and the wires have not been broken, etc. This is a cause that excludes certain defeaters. But this is an unintuitive and unsatisfactory result. Turning on the power supply is a cause of the laser firing; we do not have to specify a total cause.

A situation like this will arise anytime there is a causal defeater that renders the cause ineffective, and yet the probability of the effect is not zero or one. The above example used a laser to get an irreducible probability. It could have been done in many other ways. The important point is that a cause can be defeated in such a way that it does not raise the probability of the effect in every causally homogeneous situation. This shows that Cartwright's basic thesis, as expressed in principle CC, is wrong. Principle CC is too narrow, because it excludes genuine cases of causation.

Although principle CC turned out to be too narrow, it does not appear to be

too broad. I can think of no example in which all of the requirements are satisfied, and yet C is not a cause of E. Principle CC even seems to handle interactive forks, which is a problem for many other theories. If we have an interactive fork, with the interaction symbolized by A, and the two prongs by C and E, then it does not appear that principle CC claims that C is a cause of E. A will be a member of {C}, and in the state descriptions that contain $\neg A$, it does not seem likely that $P(E/C\&K) > P(E/K)$. Thus C is not a cause of E, which is correct. Thus until a counterexample is found, we can consider CC to give us a sufficient condition for causation.

However it is important to notice that principle CC is just expressing the conception of causation that is found in causal laws such as aspirin relieves headaches. It does not deal with specific events, and is not a generalization of specific events. The causation that CC describes deals with types of events, and is thus different from the probabilistic causation that Reichenbach and Suppes discussed.

7.3. Skyrms' Treatment of Simpson's Paradox

In a recent book, *Causal Necessity*, Brian Skyrms presents an idea of statistical invariance, which he calls resiliency. He uses the idea of resiliency to analyze ideas such as conditionals, necessity, and problems in decision theory. However, when it comes to discussing probabilistic causality, the idea of resiliency is not even mentioned. This is somewhat surprising, since the whole book is basically the task of applying the notion of resiliency to various philosophical problems. Skyrms' chapter on causation is mostly a summary of what other authors have claimed, and is not a major new work dealing with causality. However, he does discuss Simpson's paradox, and since the book has attracted considerable attention, it will profit us to discuss it.

7.3.1. The Pareto-Dominance Condition

Skyrms' briefly presents a condensed version of Cartwright's principle CC. He then gives a nice example that illustrates Simpson's paradox, and then briefly presents Reichenbach's definition of screening off. Skyrms' then presents a definition that is almost identical to Reichenbach's definition of causal relevance, but rejects this definition as being inadequate for reasons based on Simpson's paradox. Skyrms then presents a definition of probabilistic causality, which we may take to be his contribution to the discussion on probabilistic causality. Skyrms never endorses this principle, but only claims that it is an interesting weakening of Cartwright's proposal. Since we have seen that Cartwright's principle CC was too strong, perhaps Skyrms' weakening of it will be adequate. Skyrms' (1980, p. 108) definition is as follows:

Pareto-Dominance Condition

1. $\Pr(E \text{ given } C \& B_i) \geq \Pr(E \text{ given } \neg C \& B_i)$ for every B_i
2. $\Pr(E \text{ given } C \& B_i) > \Pr(E \text{ given } \neg C \& B_i)$ for some B_i in the fundamental partition.

In this condition we can consider the B_i 's to be the same as the K_j 's in Cartwright's definition. The difference between CC and the pareto-dominance condition is that whereas CC requires that the cause raise the probability of its effect in every causally homogeneous state description, the pareto-dominance condition merely requires that a cause not decrease the probability of its effect in any state description, and that it increase the probability in at least one of the state descriptions.

7.3.2. Cartwright's Rejection of the Pareto-Dominance Condition

Cartwright discusses the reasoning why she required that the cause raise the probability of the effect in every state description, and not just not decrease the probability of the effect in any of the state descriptions. Cartwright (1979, pp. 427-8) says:

Must a cause increase the probability of its effect in *every* causally fixed situation? Mightn't it do so in some, but not in all? I think not.

Whenever a cause fails to increase the probability of its effect, there must be a reason. Two kinds of reasons seem possible. The first is that the cause may be correlated with other causal factors. This kind of reason is taken account of. The second is that interaction may occur. Two causal factors are interactive if in combination they act like a single causal factor whose effects are different from at least one of the two acting separately. For example, ingesting an acid poison....

From the above passage we see that Cartwright rejects the pareto-dominance condition because she believes that it is not strict enough.

7.3.3. Principle CC and the Pareto-Dominance Condition

The pareto-dominance condition can handle some cases of causation that Cartwright's CC could not. If we take the example of the laser and the cut wires, which CC could not handle, we find that it presents no problem for the pareto-dominance condition. Principle CC had problems with this example, because in state descriptions in which the wires were cut, turning on the power source did not raise the probability of the laser firing. Thus turning on the power is not a cause of the laser firing, according to principle CC. However, the pareto-dominance condition does not require that the probability of the laser firing increase in every state description; it only requires that it not decrease in any of the state descriptions and that it increase in at least one state description. The turning on of the power supply does not decrease the probability of the laser firing in any state description, and it certainly raises it in some of them. Thus turning on the power supply is a cause of the laser firing according to the pareto-dominance condition, which is as it should be.

7.3.4. Critical Discussion of the Pareto-Dominance Condition

Problems for the pareto-dominance condition arise when we realize that there are certain state descriptions, or certain constellations of causally relevant background factors, in Skyrms' terminology, in which certain events will defeat the turning on of the power in such a way that it will actually lower the probability of the effect occurring. Suppose that the power supply was hooked up to another alternative power supply in such a manner that if the first one is turned on, then the second cannot be turned on, even if the wires between the first and the laser

are cut. In this situation the turning on of the first power supply would lower the probability that the laser will fire, which would exclude it from being a cause according to the pareto-dominance condition.

A similar situation arises in the example that I discussed concerning ingesting acid and ingesting alkali. In that example, there are some state descriptions in which ingesting acid actually lowers the probability of death; namely, those which include the ingestion of alkali. Thus the ingestion of acid is not a cause of death, according to the pareto-dominance condition. The preceding two examples show that like Cartwright's principle CC, Skyrms' pareto-dominance condition is also too strong.

A reply to these objections is available to Skyrms. Skyrms could reply that he is not interested in general laws, like Cartwright was, but rather he is interested in actual causal chains. Skyrms (1980, p. 109) expresses this application of the pareto-dominance condition in the following:

How does this analysis apply to causal factors for events? Here we want to fix the relevant factors that are in fact present. A heart attack did cause poor Cecil's death. It is true that being run over by a steamroller screens off a heart attack from death, but Cecil was not, in fact, run over by a steamroller. We can neglect those cells which include being run over by a steamroller, and indeed the coroner would like to zero in on that cell which includes the true constellation of background causal factors.

This statement of Skyrms' is an answer to the above objections. Unlike Cartwright, Skyrms is interested in actual causal chains. Because of this, he does not need to consider every possible causally relevant factor in the state descriptions, or cells, in Skyrms' terminology. Cartwright had to require that the state descriptions be formed from a set $\{C_i\}$, which was a set which contained all causally relevant factors. With certain exceptions, any factor which can either cause the effect or prevent the effect from occurring, regardless of whether that factor was present or not, was to be included in $\{C_i\}$. The reason that Cartwright could not limit the set $\{C_i\}$ to the factors that were in fact present, is that she was not dealing with actual causal chains that were made up of particular events. Thus by dealing with particular causal chains instead of general laws, Skyrms can avoid the necessity of including all causally relevant factors in $\{C_i\}$, whether they are present or not. Skyrms thus feels that he can ignore the cells or state descriptions which include events that did not happen.

By ignoring the state descriptions which contain events that did not happen, Skyrms can avoid the above problems. In the laser example, it is not true that the wires are cut and that there is an alternative power source that is connected in such a way that it cannot be turned on if the other power source is turned on. Thus it would not be a member of the B_i that is conditionalized on. Similarly, if the person did not ingest an alkali along with the acid, then ingesting an alkali would not be a member of the B_i , and thus the ingestion of acid would not lower the probability of death in this situation. We see that Skyrms can handle the above counterexamples by dealing with actual causal chains and not with general causal laws.

But if we, as Skyrms suggests, ignore all of the cells or state descriptions that contain causal factors that are not in fact present, then we will always be left with only one B_i to conditionalize on. But then Skyrms suggestion becomes just to see if the cause raises the probability of the effect in the reference class that contains all members of the causal chain up to C . We must require that the B_i be subject to the same conditions that Cartwright's state descriptions were. We require that the cause increase the probability of the effect and not just not decrease it because condition (2) of the pareto-dominance condition requires that the C raise the probability of E for some state description, and since we only have one, it must raise it in that one.

The problem with this version of the pareto-dominance condition is that it is unable to solve Simpson's paradox. We must investigate what causally relevant factors the cell B_i must contain. Cartwright did not allow $\{C_i\}$ to contain any effects of C , because otherwise principle CC would lead to absurd results. It appears as if the same stipulation must be made of B_i . But then the pareto-dominance condition, and principle CC, cannot handle many of the straightforward cases that they were designed to handle.

Let us return to Hesslow's example of contraceptives and thrombosis, which is a clear case of Simpson's paradox. The taking of oral contraceptives is correlated with pregnancy, which reverses the relevance relations. We want our cell B_i , or $\{C_i\}$, to contain the causally relevant factor of not being pregnant; when B_i and $\{C_i\}$ contain this factor, we find that the taking of oral contraceptives does raise the probability of thrombosis. However a problem arises because B_i and $\{C_i\}$

cannot contain the factor of not being pregnant, because that is an effect of taking oral contraceptives. We cannot allow effects of C into B_i or $\{C_i\}$, but often the reason that C and some other factor are correlated is because the other factor is an effect of C. It appears as if we want to exclude some effects of C from B_i and $\{C_i\}$, but not all of them. This poses a problem for the very core of these proposed solutions to Simpson's paradox. Thus until some solution to it is given, we must conclude that the pareto-dominance condition, like principle CC, is faulty.

Another type of problem that the pareto-dominance condition faces is the same problem that faced Reichenbach's definition of causal relevance. We saw that Reichenbach's definition of causal relevance could not pick out actual causal chains. In that context, we discussed an example in which two marksmen fire rifles at a target, but only one of them hits the target. Certainly the firing of the rifle that missed the target is not a cause of any event at the target, such as the balloon breaking. Suppes' theory also had problems with this example, and so does the pareto-dominance condition. The pareto-dominance condition is fulfilled for this example, and yet we do not have an actual causal chain connecting the events. Because of limitations on what can go into B_i , Skyrms cannot fill in any of the links between C and E, as Good attempts to do. Thus there is no chance that his theory will be able to handle this problem. It is beginning to appear that all probabilistic definitions of causality can only pick out possible causal chains, and not actual causal chains.

7.4. Towards A Solution of Simpson's Paradox

I believe that the difficulties brought out by Simpson's paradox are important, and threaten to undermine the essence of probabilistic causality. Although Cartwright's and Skyrms' solutions to the problem ultimately fail, perhaps we can learn from their intuitions and make a small contribution to a solution of the problem.

To begin, I will be dealing with specific causal chains, and not with general causal laws. Perhaps if we find a solution for specific causal chains we can generalize it and apply it to general laws. The reason that Simpson's paradox arises is because the cause can be correlated with a third factor in such a way that the

statistical relevance relations are reversed. It appears that the easiest way to avoid the problem is to require that there be none of the correlations that might reverse the statistical relevance relations. There seem to be two ways in which the statistical relevance relations can be reversed, and both of these must be accounted for in order to avoid Simpson's paradox. Let us consider the following principle:

** C is a *prima facie* cause of E if and only if either

1. $P(E/C) > P(E)$ and $\neg(\exists F \text{ such that } P(E/F) > P(E) \text{ and } P(E/C \& F) \leq P(E/\neg C \& F))$ or
2. $P(E/C) \leq P(E)$ and $(\exists F \text{ such that } P(E/F) \leq P(E) \text{ and } P(E/C \& F) > P(E/\neg C \& F))$.

Principle ** does not attempt to deal with the issues of spurious causation. It is designed to merely solve the problem, raised by Simpson's paradox, that a cause may not raise the probability of its effect, or that things that raise the probability of the effect are not causally connected with the effect. Although this principle does not claim to be able to pick out causes solely on the basis of probability relations, it does claim that an increase in probability in the manner specified is necessary for causation. Criterion (1) requires that a cause not raise the probability of its effect simply because it is correlated with something that does raise the probability of the effect. This condition attempts to require that C be a genuine causal factor of E. Criterion (2) accounts for cases in which the cause actually lowers the probability of the effect because it is correlated with some negative cause of the effect. In order for an event that lowers the probability of another event to be a cause of the second event it must be correlated with a negative cause of the second event.

We should realize that principle **, like principle CC, is not an analysis of causation, because the idea of causation appears on both sides of the principle. Principle ** is intended to suggest some restraints upon what constitutes a *prima facie* cause. Principle ** does not solve Simpson's paradox, because it runs into the same problems that Skyrms' solution did. We must find a way to characterize what we are going to count as a legitimate correlation, and which ones we wish to ignore. Although it is not an adequate solution to the problem, I do think that principle ** does bring out some of the intuitions we would like to capture in a solution of Simpson's paradox. Hopefully an adequate solution can be developed, and perhaps principle ** will be helpful in this project.

CHAPTER 8

SALMON'S THEORY OF CAUSALITY

All of the theories of probabilistic causality that we have discussed so far have been attempts to explicate causality and causal connections solely in terms of probability relations. We have seen that all of these theories have failed in their attempt to use probability to identify events in certain causal chains. Although most discussions of probabilistic causality assume that probability relations must pick out probabilistic causes, that idea is by no means crucial to a theory of probabilistic causality. Beginning with a series of articles, and culminating in a recent book, Salmon has attempted to develop a theory of causality that will account for cases in which the causal connection is probabilistic and not deterministic. For our purposes, the interesting aspect of his analysis is that he does not base his theory of causation on probability relations. This difference will result in a theory quite different from the preceding theories. In this chapter I will briefly present the basic features of Salmon's theory of causality. I will then critically discuss the theory, paying particular attention to the two basic principles upon which it is based.

8.1. Explication of Salmon's Theory

8.1.1. Causal Production and Causal Propagation

When Salmon considers the causal structure of the world, he finds two concepts that are fundamental to our idea of causation: causal production and causal propagation. Causal production is the means by which something produces something else. An example might be, an automobile's collision with a wall produces a wrecked automobile and a collapsed wall. Causal propagation is the means by which causal influence is transmitted through spacetime. An example might be radio waves, which propagate causal influence. Salmon believes that although these two concepts are related, we should look at each one as basic, and

not to be explained in terms of the other. A large part of Salmon's theory is an attempt to work out the details of the concepts of production and propagation.

8.1.2. Causal Processes and Pseudo-Processes

The fundamental means by which propagation of causal influence takes place is by causal processes. Accordingly, processes will play a fundamental role in Salmon's theory of causality. Salmon wants processes to be the basic ontological entities in his theory. This is different from all of the theories that we have considered so far, which construed events as the basic ontological entities. Examples of processes might be, radio waves, the shadow of a car, a beam of light, a rock, etc. Processes are different from events in that they usually last longer than events do, as well as often being more spread out in space. In the standard spacetime diagrams, processes are diagrammed by lines, and events are diagrammed by points. Although processes are fundamental to Salmon's ontology, not every process is a causal process. Some processes are pseudo-processes, which means that they are not causal processes. A pseudo-process is not capable of transmitting causal influence, whereas a genuine process is capable of transmitting causal influence. An example of a pseudo-process is a shadow. A shadow cannot transmit any causal influence, or any part of its structure; thus it is not capable of being the means of causal propagation. The distinction between processes and pseudo-processes is a crucial distinction.

8.1.3. The Transmission of Marks

Since causal processes can transmit some of their own structure, if we modify the structure, what is transmitted will also be modified. This observation will help us to distinguish causal processes from pseudo-processes. If we modify the structure of a pseudo-process, that modification is not passed on to later states of the pseudo-process. For example, a beam of white light, which is a causal process, can be modified by placing a red filter in its path. After the light beam passes through the filter, the light is red on the other side. The structure of the light beam has been modified by the filter, and this modification is passed on to later states of the process. Now consider a rotating light source, which sends out a beam

of light onto a circular enclosure surrounding the light source. There will be a spot of light that travels around the surface of the enclosure. This spot of light is a process, but it is also a pseudo-process; the spot of light cannot transmit any of its structure to a later part of the process. If we were to mark the process by placing a red filter between the light source and the wall, we would modify the structure of the light spot on the wall. But this modification could not be passed on to another spot on the wall. When the light source rotates, the spot will return to its original color, unless it is also marked. From these two examples we see that causal processes transmit their structure, whereas pseudo-processes are unable to transmit any of their structure or any type of information.

Although the intuitive idea of mark transmission has been presented, Salmon must still give a more precise analysis of it. One might question what it means for causal structure or a mark to be transmitted to a later part of a causal process. Salmon (forthcoming, p. 165) presents the following principle, which is designed to be an analysis of mark transmission:

MT Let P be a process which, in the absence of interactions with other processes, remains uniform with respect to a characteristic Q which it manifests consistently over an interval which properly includes both of the spacetime points A and B . Then, a mark (consisting of a modification of Q into Q'), which has been introduced into process P by means of a single local interaction at point A , is *transmitted* to point B if P manifest the modification Q' at B and at all stages of the process between A and B without additional interventions.

This principle gives us a criterion of mark transmission. The main idea, based on Russell's at-at theory of motion, is that for a mark to be transmitted from A to B is for the mark to appear at A and B , and at every point in between.

Once Salmon has defined what it is for a mark to be transmitted, he can give a principle of structure transmission:

ST If a process is capable of transmitting changes in structure due to marking interactions, then that process can be said to transmit its own structure (forthcoming, p. 174).

This principle tells us that if a process can transmit a mark, then it transmits its own structure. The ability to transmit its own structure is important because it is a characteristic of causal processes, and not of pseudo-processes. Salmon (forthcoming, p. 175) then gives a principle of propagation of causal influence,

which makes explicit the idea that if a process is capable of transmitting its own structure, then it is capable of propagating causal influence:

PCI A process which transmits its own structure is capable of propagating a causal influence from one space-time locale to another.

This principle tells us that if a process can transmit its own structure, then it can transmit causal influence. Principles MT, ST, and PCI give us an account of what it means for a process to transmit causal influence through spacetime. These principles form the basis of Salmon's theory of propagation.

8.1.4. The Principle of Causal Interaction

We must now consider causal production. Our preceding discussion of causal propagation actually assumed that we could make sense of causal production, because the marking process is a case of causal production. Salmon basically thinks that causal production is a result of an interaction between two or more causal processes. Causal processes may interact when they intersect one another. Salmon (forthcoming, p. 192) presents a principle of causal interaction to tell us which intersections of processes count as causal interactions:

CI Let P_1 and P_2 be two processes which intersect with one another at the spacetime point S which belongs to the histories of both. Let Q be a characteristic which process P_1 would exhibit throughout an interval (which includes subintervals on both sides of S in the history of P_1) if the intersection with P_2 did not occur; let R be a characteristic which process P_2 would exhibit throughout an interval (which includes subintervals on both sides of S in the history of P_2) if the intersection with P_1 did not occur. Then, the intersection of P_1 and P_2 at S constitutes a *causal interaction* if: P_1 exhibits the characteristic Q before S, but it exhibits a modified characteristic Q' throughout an interval following S; and P_2 exhibits the characteristic R before S, but it exhibits a modified characteristic R' throughout an interval following S.

This principle gives a criterion for causal interactions. Not every intersection of processes is a causal interaction; one or more of the processes could be a pseudo-process. It is even possible for two causal processes to intersect and not interact. If two light beams intersect each other, it is highly probable that a causal interaction will not take place. Since causal production is due to causal interaction, with this principle Salmon has given us an account of causal production.

We can now see what Salmon's theory of probabilistic causality would be like. Two events are causally connected if they are linked by a causal process. It does not matter if those two types of events are normally not linked by a causal process, or whether they normally are linked by the causal process. What is important is that there is a causal process connecting these two events. The probabilities involved do not affect whether the events are causally connected. We thus see that Salmon's theory is not attempting to use probabilities to determine when two events are causally connected, but rather he gives criterion independent of probabilities to determine if two events are causally connected. In probabilistic causality, Salmon thinks that the important aspect is the causal process connecting two events, and not the increase in probability. This may be the reason why Salmon was one of the first to question whether causes must really be positively relevant to their effects.

8.2. Critical Analysis of Salmon's Theory

8.2.1. The Counterfactual Aspect of Principle MT

At this point we shall take a closer look at some of the central points of Salmon's theory. We immediately notice that both MT and CI are counterfactuals. The counterfactual element of MT arises in the first sentence, where it is claimed that the process remains uniform with respect to Q, in the absence of interactions with other processes. The idea is that if the process P would have manifested characteristic Q between A and B, but didn't, because Q was changed into Q', then a mark has been introduced. This basically claims that the process would not have had the mark had it not been introduced at A.

However, principle MT is ambiguous. We could interpret principle MT as requiring that either one of the following counterfactuals be satisfied:

1. If the process P were marked at point A by changing characteristic Q into Q', then P would have Q' at B and all stages of P between A and B,
2. If the process P had not been marked at point A by changing characteristic Q into Q', then P would have had Q at B and all stages of P between A and B.

These two counterfactuals are not the same counterfactual, nor are they equivalent;

either could be true when the other is not true. When discussing possible counterexamples to principle MT, we must remember these two possible interpretations.

Salmon requires principle MT be counterfactual in order to handle an objection brought forth by Nancy Cartwright (Salmon forthcoming, p. 165). Cartwright's objection is that the process does not qualify as marked if the mark would have been there even if the mark was not introduced. Consider the example of the rotating light source. Suppose that a short time before the spot was marked with a red filter at the wall, a red lens was placed over the light source so that the spot on the wall will remain red even after it passes the spot where the filter is. The process is marked at the spot on the wall, and it appears that the mark continues with the process. This is an undesirable situation, because the spot on the wall is a pseudo-process and cannot transmit a mark. The counterfactual formulation of MT is designed to eliminate counterexamples such as this. According to MT, the mark is not transmitted along the wall because the mark would have been there even if it had not been introduced at point A. This reasoning appears to make use of the second interpretation of MT. It is false that if the process had not been marked, that the mark would not appear in the process. Although it is unfortunate that it was necessary to appeal to counterfactuals to formulate the principle of mark transmission, Salmon notes that the counterfactual involved is easily testable.

8.2.2. Problems With Principle MT

Although principle MT handles Cartwright's counterexample, there may be similar counterexamples that are more problematic. Suppose that we change Cartwright's example slightly and claim that placing the red lens on the light source is an effect of marking the spot on the wall with the red filter. This would be possible if someone was trying to make the spot on the wall look like a genuine process and not a pseudo-process. In this case the same counterfactual is satisfied: if the process had not been marked at point A (where the filter is placed on the wall), then the mark would not have appeared later in the process. The causal connection between marking the process at point A and placing the lens on the light source supports the counterfactual in MT. This example would tend to

indicate that MT is too broad: according to it some pseudo-processes can transmit marks.

A closer examination of MT shows that the previous example may not be a counterexample to MT. Principle MT requires that P be marked by a single local interaction at A. In the above example it will take a certain amount of time for a causal signal to travel from the marking process at A, to the light source, and then back to the wall. Since there is an upper limit on the speed with which causal signals can be sent, there is a certain time interval from the moment the spot on the wall is marked until the mark will appear on the wall at a later part of the process. During this interval, the process will not be marked, and thus principle MT will not claim that a mark is transmitted through the process, since the mark does not appear at every place in the interval between A and B.

In response to the above objection to the presented counterexample, there seem to be two responses. The first response is to change the example slightly and make the red filter that marks the spot on the wall be rather large. If this were so, and if the light source were not rotating very rapidly, then by the time the spot on the wall was not red due to the red filter, it would be red due to the red lens at the light source. In this example, the spot of light on the wall is red at every point after it is marked, and it seems to be true that if it had not been marked, then it would not have been red. This problem arises because the mark is continually introduced until the light arrives red from the light source. We will discuss this problem in more detail when it arises in terms of a more general problem.

A second response to the defense of MT would be to change the example and make the marking of the process at A and the placing of the red lens at the light source both effects of a common cause. Suppose that some common cause causes both the red filter to mark the spot on the wall at A, and slightly earlier it caused a red lens to be placed in front of the light source. In this case, there would be no place on the wall where the spot is not red after A; after the mark is introduced, it remains continuously. The question that arises here is whether the situation we have set up supports the counterfactual that if the spot hadn't been marked at A, then the spot would not have been red after A. I believe that we can set the situation up in such a way that the common cause is necessary and

sufficient in the situation for each of its effects, in which case it would seem to support the counterfactual. Suppose that someone were trying to construct a counterexample to principle MT, and in order to do so he decides to place the red lens over the light source if and only if he will place the filter on the wall slightly later. Then it seems to be true that if he hadn't placed the filter at the wall, that he wouldn't have placed the lens on the light source, and the process would not have been marked. In this situation it appears that MT would wrongly classify the revolving spot on the wall as a causal process, since the red mark would be transmitted.

The above examples are based upon a fundamental problem with the idea of marking. A causal process transmits a mark because it transmits its own structure. We can make a pseudo-process appear to transmit a mark by many interactions of causal processes. In the above examples, the revolving spot of light, which is a pseudo-process, appeared to transmit a mark because there were many causal interactions occurring. When the red light from the light source interacted with the wall, a red spot appeared. The red spot appeared, not because a previous part of the process was red, but rather because light of a certain frequency interacted with the wall. Consider another example. Suppose we have a car moving along the road, and that we dent the car in such a manner that the shadow of the car is also modified. We have marked the shadow, by marking the car. Unfortunately, the shadow remains marked after the initial marking, and it appears that if it was not marked when it was that the mark would not be a part of the process. This is because the car was marked, and there is a strong correlation between the car and its shadow. According to MT, it would seem that the shadow, which is a pseudo-process, is capable of transmitting a mark. But as before, the mark does not appear at later stages in the process as a result of an earlier stage in the process, but rather because it is continually being marked. By continually marking a pseudo-process it is possible to make it appear as if the pseudo-process is transmitting a mark.

One might object to this counterexample by claiming that the shadow is not marked by a single *local* interaction. We could remedy this difficulty by noticing that there are many similar examples in which the interaction would be local. Suppose we have a piece of paper lying on a table. The shadow under the paper

will be marked if we tear some of the paper. In this example, the paper is very close to the shadow, so tearing it will count as a local interaction. Thus we cannot save principle MT from this counterexample by an appeal to locality.

The above examples rely on the claim that the counterfactual condition of MT is satisfied. That is not an obvious fact. In the example where the filter and lens are a result of a common cause, one might want to claim that if the filter were not placed at the wall, then the lens might still be placed at the light source, and the mark might appear in the process. In this case we could only claim that the process might not have had the mark after the interaction; we could not claim that it would not have it. This interpretation of the counterfactual does have some appeal to it, although it is certainly controversial.

However it seems that the example of the car and the shadow still supports the counterfactual requirement of MT. If the shadow was not modified at point A, it seems to me that the car would not have been dented then either, and thus the car and shadow would not be modified at a slightly later time. Unlike the previous case, there doesn't seem to be any way that the shadow might not be modified and the car still dented in the same manner. Of course, the light rays could bend in some miraculous fashion, but I consider that irrelevant to this issue. It seems that if the shadow were not modified, neither would the car have been dented. Since I do not find it plausible to deny that this example satisfies the counterfactual requirement of MT, we must find some other way to exclude these examples.

There is also a less serious problem with MT that needs to be noted. Suppose that we have a process P and we let Q be the property of not having been marked at A, and Q' be the property of having been marked at A. In this case, if the process was not marked at A, it will have property Q throughout, but if it was marked at A, it will have property Q' throughout. This will be true regardless of whether P is a causal process or a pseudo-process. Accordingly, all processes are capable of transmitting a mark, according to MT. There will always be a Q' that both a pseudo-process and a causal process can transmit.

The problem here arises because of the nature of Q and Q'. We are not able to tell if a point in a process has property Q or Q' without looking at other parts of the process. Reichenbach (1956, p. 187) would say that the classes defined by the properties Q and Q' are not codefined. If we require that the classes

defined by Q and Q' be codefined, this problem disappears. The class of events in a process that was or was not marked at A is not codefined.

8.2.3. Principle CI and Pseudo-Processes

When two processes intersect, Salmon thinks that there can be a causal interaction only if both of the processes are causal processes. Two pseudo-processes cannot causally interact, nor can a causal process and a pseudo-process interact. When examining principle CI, we notice that it requires that both processes are modified after the intersection in order for there to be a causal interaction. If Salmon did not require that both processes be modified, it would be possible for a causal process to causally interact with a pseudo-process. Suppose we have a spot of light moving along the wall, and that it crosses some unexposed photographic film. The spot of light modifies the photographic film in some way, but after the spot of light passes the photographic film, the spot of light is unchanged. This does not count as a causal interaction according to CI, because the spot is unchanged after the intersection. If we only required that one of the processes be changed after the intersection, this intersection would count as a causal interaction, which is undesirable. Of course, there is a causal interaction that takes place, but it is not an interaction between the spot on the wall and the photographic film; on the contrary, it is an interaction between the beam of light and the photographic film.

Although there may appear to be interactions between two causal processes in which one of the processes is not modified, upon closer examination we find that they are both modified. When we place a red filter in the path of a beam of light, it appears that the light is modified, but it is harder to see the modification to the filter. Upon close inspection, we do find that the filter is modified and increases in energy (Salmon forthcoming, pp. 191-2). In order to apply CI in every interaction, it may be necessary to specify the processes very precisely in order to notice the modification of both processes.

8.2.4. Problems With Principle CI

Let us now return to our example of the rotating spot and the photographic film. However, now suppose that there is an observer who places a red lens in front of the light source if and only if the spot and photographic film are going to intersect each other. This has the result that if they do intersect, the spot is red after the intersection. In this situation, it appears that principle CI is satisfied, and the intersection of the spot and the photographic film is a causal interaction. It seems that the counterfactual is satisfied, because if the spot and photographic film did not intersect, the spot would be white instead of red and the photographic film would also be different. This example shows that sometimes CI will classify an intersection between a causal process and a pseudo-process as a causal interaction.

This example is similar to the ones discussed in connection with principle MT. Thus it too relies on the fact that if the spot and photographic film are not going to intersect, the lens will not be placed in front of the light source. Whether or not this is true, depends upon what factors we hold fixed in evaluating the counterfactual situation. I have attempted to describe a situation in which it seems likely that the counterfactual requirement of CI is satisfied. However, I will grant that this is a questionable example, and one could deny that it is a counterexample to principle CI.

Although it is possible to claim that the counterfactual requirement is not satisfied in the above example, I do believe that it is possible to construct other counterexamples in which it is evident that the counterfactual requirement is satisfied. These examples will be similar to the shadow and car example that we discussed in connection with principle MT. The following example shows that according to CI, two pseudo-processes can intersect and be classified as a causal interaction. Suppose we have two billiard balls that are rolling on a glass surface. Above the surface, a light shines on the glass in such a manner that shadows of the balls are projected on a second surface below the glass. Suppose that the two balls approach each other, collide, and bounce off in different directions. This is a perfectly legitimate case of a causal interaction. Both of the billiard balls are processes, and each of the processes is modified by the intersection of the two processes. Now consider the shadows of the balls on the surface underneath the glass. The shadows of the balls are processes, and it appears that they are also

modified by the intersection of the two shadows. If the shadows had not intersected, it seems as if they would have gone in different directions than they did, which is what is required by CI. The counterfactual requirement of CI is satisfied, and thus according to CI, the intersection of the shadows constitutes a causal interaction. But since the shadows are merely pseudo-processes, their intersection cannot be a causal interaction.

This difficulty is also similar to the ones that arose for principle MT. The reason the two pseudo-processes support the required counterfactual and appear to have a causal interaction is because they are dependent upon two processes which support a counterfactual and have a causal interaction. If the two billiard balls, which are causal processes, did not causally interact, neither would the shadows have the appearance of a causal interaction. Principle CI needs to exclude from the class of causal interactions intersections of processes that support counterfactuals simply because they are dependent on other causal processes that do support the counterfactuals.

8.2.5. The Relationship Between Principle MT and Principle CI

The above counterexamples to principle CI all involve an intersection of processes in which at least one of the processes is a pseudo-process and not a causal process. The natural solution to this problem would be to simply require that the processes be causal processes and not pseudo-processes, in principle CI. If we require that the processes that intersect be causal processes, the above counterexamples will not arise.

The problem with this solution to the problem is that it would render Salmon's theory circular. Salmon would be guilty of defining MT and CI in terms of one another. Principle MT was given in order to distinguish causal processes from pseudo-processes; in order to distinguish among them, a marking technique is employed. But in order to mark a process, there must be a causal interaction. If two processes intersect, but do not causally interact, a mark will not be made. Also, principle MT assumes that process P does not undergo any additional causal interventions between A and B; but this requirement in MT presupposes that we have an account of what constitutes a causal interaction. Thus in order to define principle MT, we need an adequate account of causal interactions. Because principle

MT presupposes an adequate account of causal interactions has been developed, we cannot use MT to restrict the processes that principle CI applies to. This precludes us from restricting CI to causal processes.

In order to mark a process, whether it be a causal process or a pseudo-process, a causal interaction must occur. One may question how it is possible to mark a pseudo-process, since a pseudo-process cannot causally interact with other processes. The answer to this lies in the fact that pseudo-processes are dependent upon causal processes, and causal processes are capable of interaction. Consider the rotating spot on the wall. At the spot, we have an intersection of three processes: the spot, the wall, and the beam of light. Since both the wall and the beam of light are causal processes, they are capable of causal interactions. The causal interaction will take place at the same location as the intersection with the spot of light, because the spot of light is a result of the interaction between the beam and the wall. By marking one of the causal processes that intersect here, we mark the spot on the wall. This fact also illustrates the point that we must have an analysis of causal interactions in order to give an adequate analysis of mark transmission. Thus we cannot use the definition of mark transmission to restrict the definition of causal interaction to causal processes.

This points out how intimate the relation is between principles MT and CI. The problems that arose for both of these principles is basically the same problem; a pseudo-process can behave like a causal process if it is dependent upon an appropriate causal process. By being connected with causal processes in appropriate ways, pseudo-processes can appear to transmit marks or causally interact with other processes. It appears that if we could adequately define one of these notions, we could use it to define the other: if we could define mark transmission without causal interactions, we could use it to define causal interactions, and if we could define causal interactions, we could use it to define mark transmission. For example, if we had an adequate definition of causal interactions, we could define a causal process as a process which is able to causally interact with other processes. A pseudo-process is a process that is not capable of interacting causally with other processes. The reason that the counterexamples to MT and CI appear similar is due to the intimate relation between MT and CI.

8.3. Different Types of Causal Interactions

In addition to the types of interactions in which two processes intersect, and then continue separately, Salmon (forthcoming, pp. 203–4) discusses two other types of causal interactions. A λ type interaction is where two processes converge into a single process, and a Y type interaction is where a single process separates into two separate processes. The type of interactions that were characterized by CI are called X type interactions. Salmon expressed a desire to see how to explicate causal interactions in terms of Y and λ type interactions, which appear to be simpler types of interactions than the X type.

I had originally thought that Y and λ type interactions could be adequately characterized by the various properties of mark transmission. In a λ type interaction, if a mark is capable of being transmitted from the two earlier processes into the final process individually, then it appears that a causal interaction has occurred. For example, suppose that we have two balls of putty that collide and stick together; this is a λ type interaction. If we mark either one of the balls, the mark will appear in the final ball of putty. Either prong is capable of transmitting a mark to the final process. Similar reasoning will apply to Y type interactions. In a Y type interaction, it should be possible to transmit a mark from the single process to each of the two processes that it separates into. I am not claiming that the same mark must be able to be transmitted, but only that some mark be capable of being transmitted. This seems to be true in every case of Y and λ type interactions.

The problem with this explication of causal interactions is that it would be circular. As we saw earlier, we cannot use the principle of mark transmission to define causal interactions without becoming involved in circular reasoning. The above claims are true, but they are unsuitable for use in a definition of causal interaction.

One might claim that if we read principle CI somewhat liberally, that it already accounts for Y and λ type interactions. Let us first consider λ type interactions. Suppose that P_1 and P_2 are the two processes that intersect and form one process. Since principle CI does not specifically require that the two processes be different processes, we will assume that P_1 and P_2 are the same process after the interaction. In this case, there is a property Q which process P_1 would have had if

the intersection with P_2 had not occurred: Q would be any property that would have been different, such as being a different process from P_2 . Similarly, there is a property R that P_2 would have had if P_2 did not intersect P_1 . Perhaps this property is being a separate process from P_1 . If P_1 and P_2 do causally interact, there must be properties Q and R that fit the requirements. After the intersection, processes P_1 and P_2 have properties that they would not have had if they had not intersected. This property may be as simple as being the same process. Principle CI does not seem to rule out the possibility of P_1 and P_2 becoming the same process. Thus we see that λ type interactions are handled by principle CI. Similar considerations will show that Y type interactions are handled by CI also.

The problem with this interpretation of principle CI is that it would allow the intersection of pseudo-processes to be classified as λ and Y type interactions. Suppose that we have a blue spot of light moving on the wall which intersects with a yellow spot of light; the result will be a green spot of light. The blue and yellow spots of light are pseudo-processes, but they join together to form another pseudo-process, the green spot of light. This intersection of pseudo-processes would be classified as a genuine causal interaction according to the liberal interpretation of principle CI. Thus in order to exclude counterexamples such as this one, we must interpret CI as applying to only X type interactions. An adequate characterization of λ and Y type interactions has yet to be found.

8.4. Interactive Forks and Causal Screening Off

The reader will recall that in conjunctive forks the common cause screens off each of the prongs from one another. In contrast to this, in interactive forks the common interaction does not screen off each prong from the other. In fact, usually one of the prongs will screen off the common cause from the other prong. But although the statistical screening off fails in interactive forks, Salmon (forthcoming, p. 196) does note that there is a sort of causal screening off that holds for interactive forks. Suppose that C is the causal interaction between two processes, and that A and B are events in the two prongs of the two processes. Salmon notes that if we mark one of the processes between A and C , that mark will not show up in the process between B and C . There is a form of causal

screening off in effect here, because one of the processes cannot affect the other process. Thus even though statistical screening off fails for interactive forks, there is a sort of causal screening off that is applicable.

Although causal screening off seems to hold in all interactive forks that involve macroscopic phenomena, there do seem to be problems with it when we consider the quantum domain. If we have two particles, which are joined in an appropriate manner and then separated, it appears as if causal screening off fails. Measuring the spin on one of the particles places it in a certain eigenstate, which also places the other particle in a certain eigenstate. Thus a mark placed on one of the processes affects the other process, and causal screening off fails.

This problem only arises in the quantum domain, and it is not obvious what should be done about it. The possible breakdown of locality in quantum mechanics is a serious problem, and few philosophers have had much to say about it. For now, we will just notice that causal screening off must be restricted to macroscopic objects.

8.5. Salmon and Causal Connections

Salmon's theory is a theory about the causal connections in the world. We might view Salmon's theory as an extension of Reichenbach's theory, because Reichenbach was also interested in the causal net of the world. Because Salmon is concerned with causal connections, it would be futile to attempt to give counterexamples to his theory on the basis that it does not coincide with the way we normally speak of causation. Salmon is not interested in picking out causes, but he is interested in picking out causal connections.

Salmon's main goal in developing a theory of probabilistic causation was to account for the role of causality in scientific explanation. Unlike other philosophers, he saw that what was important for scientific explanation was the causal connections, and not the actual probability values. Since Salmon does not think that it is possible to pick out the causal connections of the world using only probability relations, it is natural that actual probability relations play a small role in his theory. For further discussion of the role of probability relations in Salmon's theory see Rogers (1981).

APPENDIX A

INTERPRETATIONS OF PROBABILITY

When developing a theory of probabilistic causality, the nature of probability seems to play an important part of the theory. We have seen that most of the people writing about probabilistic causality believe that probability relations should be used to pick out causal relations. It appears to me that one's interpretation of probability will be relevant to whether this is possible. There are many different interpretations of probability, and it is not obvious that all of them are suited for picking out causal relations. Suppes meant for his theory to be flexible enough to be applicable to whatever interpretation of probability that one prefers. Although he does not explicitly state it, it seems that this implies that it doesn't really make any difference which interpretation of probability that we use in order to pick out causal relations. In contrast to this idea, I think that there are philosophical reasons that bear upon whether probability relations can pick out causal relations. Certain interpretations are preferable to other interpretations, for philosophical reasons, when our goal is to pick out causal relationships. In the following discussion, we will be concerned only with the features of a particular interpretation of probability that bear upon its suitability to picking out causal relations; we will ignore any other difficulties that it may face. The following criticisms are for the most part modifications of criticisms of various theories of probability that are found in Salmon (1966). For further detail on any of the specific interpretations, it should be consulted.

A.1. The Classical Interpretation of Probability

The classical theory of probability defines probability in terms of the ratio of positive to equally possible outcomes. A problem arises when we ask how to determine what the equally possible outcomes are, but this is solved by appeal to the principle of indifference. The principle of indifference states that two outcomes are equally possible or equally probable if there is no reason to think that

one is more probable than the other. If we are indifferent to two outcomes, they are equally probable according to the principle of indifference.

In dealing with probabilistic causality, the principle of indifference becomes very suspicious. There seems to be no reason to think that our indifference about two events will bear upon whether they are causally connected. The principle of indifference attempts to bring a subjective element into causation. It seems intuitively wrong to say that the causal structure of the world is dependent upon our ignorance of and indifference between events.

Another problem that arises for the classical interpretation is the one that we discussed earlier in connection with Hume's theory of probabilistic causality. Hume noticed that one out of twenty ships that leave port do not return safely. Thus, according to the classical interpretation, in this case there are twenty equally possible outcomes; in nineteen of these the ship returns safely, and in one of them the ship perishes. As I mentioned in chapter 2, it seems that the only way that one could possibly know what the twenty equally probable outcomes are is by knowledge of prior frequencies. It seems as if the classical interpretation bases its probability on frequencies when these are known, and when they are not known probability is based on ignorance. This is a strange mixture upon which to base probability. Likewise, it would be strange to say that when we have knowledge of frequencies, causation is determined by them, but if we don't happen to know what certain frequencies are, then causation is determined by our ignorance. This would be an unacceptable account of causation.

Another problem with the classical interpretation is that it can assign two different probability values to the same event. The Bertrand Paradox shows that the classical interpretation of probability does not satisfy the probability calculus, because of non-unique probability assignments. With respect to probabilistic causality, this shows that one event may be the cause of another event, and not the cause of the same event, because of two different probability assignments. But clearly the one event either is or is not a cause of the other event: it is not both a cause and not a cause of the other event. Thus we must conclude that the classical interpretation of probability is unsuitable for picking out the causal structure of the world.

A.2. The Subjective Interpretation of Probability

The subjective interpretation of probability defines probability in terms of actual degrees of belief. The probability of a proposition is the degree to which people believe that it is true. The problem with this is that often our degrees of belief do not conform to the probability calculus. Also, there seems to be no reason to think that causal relationships are determined by people's degrees of belief; people can be mistaken about causal relationships. Thus we must conclude that the subjective interpretation of probability is inadequate for the purposes of probabilistic causality.

A.3. The Personalist Interpretation of Probability

In response to these objections a modified version of the subjective interpretation has been developed. This modified version, called the personalist or Bayesian interpretation of probability, defines probability in terms of degree of rational belief. By restricting probability to degree of rational belief, these subjectivists seek to exclude probability assignments that violate the probability calculus. Thus one of the requirements of rational belief is that it be coherent, which means that it is consistent with the probability calculus. Another requirement is that one's beliefs change through time in connection with some specified rule, such as Bayes' theorem. Many, perhaps most, subjectivists think that these requirements are necessary and sufficient for rational belief.

Although the above modification does save the subjectivist from incoherent probability assignments, it still appears as if probability is far too subjective to be used as a means of determining the causal structure of the world. There are many different probability assignments that are consistent with the probability calculus. The requirement of coherence does not seem to limit the possible assignments of probability values to a single assignment, or even a small set of assignments. Anything is possible, as long as one does not break any of the theorems of the probability calculus. But clearly this is unsuitable for picking out causal relations. Causal relations are not dependent upon a person's degree of belief, even if that person's beliefs are coherent. Adopting this interpretation of probability would have

the result that causal relations would be very different for different people, since they can have very different probability assignments. This is unacceptable, because causal relations do not seem to be indexed to different people.

A personalist might reply that as long as the different people start with an open mind and change their beliefs according to Bayes' theorem, that eventually their probability assignments will become equal. Thus ultimately everyone would agree on the causal relations in the world. I think that there are problems with this, such as the time in which it might take for the probabilities to converge, but even if we ignore such problems, and problems with using Bayes' theorem as a means of changing our beliefs, it seems as if a fundamental philosophical problem remains. There seems to be absolutely no connection between causality and anyone's belief about causality, and thus there would seem to be no connection between causality and probability, if probability is degree of rational belief. The causal relations of the world just do not seem to be based upon any rational beliefs about them.

The main problem facing the subjectivist is that his probabilities are subjective, and causal relations are objective. A subjectivist might try to bridge this gap by attempting to link his subjective probabilities with objective probabilities or frequencies. One way to do this is known as the frequency principle. The frequency principle states that if the frequency of a certain event or proposition being true is known, then one's degree of rational belief must also be the same. For example, if it is known that 50% of coin flips turn up heads, then a personalist should believe that this coin will come up heads to a degree of .5. The frequency principle is an attempt to ensure that subjective probabilities are the same as objective probabilities, when the objective probabilities are known.

Leaving aside any problems with this particular version of the frequency principle, it still seems as if this modification will not solve all of the preceding difficulties. Presumably causal relations exist independently of our knowledge of them or the frequencies with which events occur. If we are ignorant of the frequency with which a certain event occurs, it seems unjustified to claim that our degrees of belief about the event define some of its causal relations. Causal relations are independent of our degrees of belief even when we are ignorant of certain frequencies.

We must conclude that the subjectivist interpretation of probability is inadequate for the purposes of probabilistic causality. There just does not seem to be any connection between degrees of belief, and causal relations. Even when the frequency principle is added, the probabilities do not become objective enough to function in a theory of probabilistic causality. One possible way for the subjectivist to get probabilities that are objective enough for probabilistic causality would be to adopt the following modification of the frequency principle:

If the frequency of a certain event or proposition being true exists, then a rational person's degree of belief must equal it.

This modification would give objective probabilities, and solve the problems that the subjectivist interpretation faces. The problem with this modification is that it undercuts the motivation for accepting the subjectivist interpretation of probability in the first place. The reason people are attracted to the subjectivist interpretation is that it gives a means of determining probabilities even when we do not seem to have knowledge of frequencies. The above principle prevents a subjectivist from doing that, and makes his theory very close to a frequency interpretation of probability. We thus see that unless radical modifications are made to the subjectivist interpretation of probability that it will not be able to play much of a role in a theory of probabilistic causality.

A.4. The Logical Interpretation of Probability

The leading proponent of the logical interpretation of probability is Rudolf Carnap. According to the logical interpretation of probability, probability is a logical relation between evidence and hypothesis. Since probability relations are logical relations, they are analytic if they are true, and inconsistent if false. Probability statements are not contingent; they are necessarily true or necessarily false.

Carnap believes that inductive logic, like deductive logic, is a semantical concept and is intimately connected with a language. Suppose we have a first order language with a finite number of one-place predicates. Within this language there will be several complete descriptions of the way things could be; these complete descriptions are called "state descriptions." We can then define the range of a

statement to be the set of state descriptions with which it is logically compatible. Since both scientific hypotheses and evidence statements are linguistic entities, they will both have a range; hypotheses and evidence statements will be compatible with some and incompatible with other state descriptions. We will then define the degree of confirmation to be a relation between the ranges of the hypothesis and the evidence. Alternatively, we could have defined degree of confirmation to be a relation between the possible worlds that the hypothesis and evidence are consistent with.

Although it seems clear that on the logical interpretation probability should be a relation between the ranges of the hypothesis and evidence, exactly what relation it should be is far from clear. One must determine how much weight to assign to each state description or possible world; perhaps some state descriptions are more important than others. Carnap calls the functions that assign weights to each state description "measure functions." We will then define the degree of confirmation of h on e , $c(h,e)$, corresponding to each measure function M by $c(h,e)=M(h\&e)/M(e)$. Thus the degree of confirmation of h on e will be a ratio of the range of the evidence and the hypothesis to the range of the evidence. There are an infinite number of different measure functions to choose from, and deciding which one of them to use is probably the biggest problem facing the logical theory of probability.

We might be tempted to assign each state description equal weight, as Wittgenstein did. However, the measure function that assigns each state description equal weight, called M^+ , is unacceptable. M^+ is unacceptable because the frequency with which an event occurs in conjunction with another event would be totally irrelevant to whether they are causally connected. But it certainly seems plausible that if two events are causally connected that they occur together more often than we would expect if they were independent of each other. Since M^+ does not modify probabilities according to how often events occur, causes would be picked out a priori if we were to accept it as our measure function. Thus in our discussion of the applicability of the logical theory of probability to probabilistic causality we will assume that some measure function other than M^+ has been picked out.

Upon the logical interpretation of probability we have seen that the statement $c(h,e)=r$ is analytic if true, and contradictory if false; this feature of the logical

interpretation raises problems in connection with probabilistic causality. If $c(h,e)=r$ is analytic, then it is necessarily true and true in every possible world. Suppose we have two events, A and B, such that $c(B,A)>c(B)$. Since probability statements are analytic, $c(B,A)>c(B)$ will either be necessarily true or necessarily false. Thus, assuming nothing screens off A from B, A will either necessarily be a cause of B, or necessarily not be a cause of B. Even in situations in which A and B are not connected or correlated, A will be a cause of B if $c(B,A)>c(B)$. What this entails is that all causal relations would be the same in all possible worlds, since they are based upon necessary truths. If A is a cause of B in one situation, it would be a cause of B in all situations. Clearly this is unacceptable for a theory of causation.

One might reply to this criticism by claiming that the above is true when we know nothing except that the two events occur, but that usually we are in a situation to know much more. It might be claimed that if we require that all of the relevant evidence be included in the calculation of the probabilities, then the probability relations will vary across possible worlds. Let i designate all of the relevant background knowledge; the central requirement of probabilistic causality then becomes $c(B,A\&i)>c(B,i)$. With this requirement of total evidence the above statement, although analytic, becomes indexed a particular possible world or set of worlds, because the background evidence is only true in some worlds. Thus we are not in the situation in which all causal relations are necessary and the same in all possible worlds. As long as there is some background information, the probability relations will differ. If we have no relevant background information, then all causal relations are necessary and the same in all possible worlds.

If we have no relevant background information, the logical interpretation of probability would pick out causes a priori. Since causation is a fact about the world, we would be able to know some empirical facts about the world a priori if this were possible. Causal claims would be synthetic a priori, when there was no background evidence. This strikes me as being implausible, and shows a flaw in using the logical interpretation to determine causal relations.

Even if we have relevant background information, one can question the appropriateness of allowing some a priori factors to influence the probability relations. This problem is particularly noticeable when confronted with different measure functions. Even if we ignore the non-standard measure functions, a

philosophical problem remains. Why should any a priori elements be relevant in picking out causal relations? There seems to be no justification at all for letting any a priori factors influence what is a cause or not, unless they are somehow about the world, which is to say that they are synthetic a priori. No matter how much background information we have, the logical interpretation of probability modifies it according to certain a priori standards. It is this modification that is problematic.

We must conclude that the logical interpretation of probability is inadequate for a theory of probabilistic causality. The causal relations of the world must be discovered empirically. Thus any theory that claims to be able to know the causal structure of the world a priori must fail. In this context it is profitable to remember Hume's distinction between matters of fact, and relations of ideas. The logical theory of probability cannot pick out causal relations unless that distinction is forgotten.

A.5. The Propensity Interpretation of Probability

The propensity interpretation of probability seems well suited to most cases of probabilistic causality. It is objective, and is not based upon any a priori beliefs about the world. But most importantly, it defines probability in terms of the causal dispositions that we are interested in. In probabilistic causality, we are interested in probabilistic causal dispositions; it seems likely that these lie at the core of the idea of probabilistic causality. But the propensity theory of probability simply defines probability in terms of these dispositions, so it seems obvious that the propensity interpretation is adequate for the purposes of a probabilistic theory of causality.

Although it seems to be true that most cases of probabilistic causation are based upon probabilistic causal dispositions, it may be the case that not all of them are. Ian Hacking (1980) has recently brought up the possibility that not all probabilities may be grounded from below. In other words, perhaps there are probabilities that properly refer only to classes and not to individuals. The question is whether the frequencies that exist are always due to some probabilities about individuals, or whether sometimes the frequencies apply only to classes and are not based upon any probabilistic facts about the individuals in the class.

Hacking (1980, p. 114) presents an example which is supposed to illustrate the fact that there may be probabilities that are not grounded from below. The example concerns Germany's falling birth rate:

Knodel found that in the national decline in birth rate is nicely mirrored in each administrative district or *Kreis*. The numbers are as regular as any which could be hoped for in demography. But when we pass to smaller units, such as the village, the uniformity collapses. Although every *Kreis* is doing the same as every other, villages within *Kreise* are all going their own ways, without much in the way of underlying laws. Now suppose that the probability facts about the *Kreis* are to be grounded from below, in probabilistic facts about individual married couples. What facts? In the reference class consisting of the *Kreis* we find homogeneity, but the probability of having say 2 children, relative to that class, is certainly not what one would infer from the villages within the class, nor any other plausible social segmentation of the *Kreis*. It may turn out that there is simply no quantitative propensity to have 2 children that can realistically be ascribed to individual couples, although there is a probability fact about the population at large.

Hacking continues and discusses the fecundability of couples, which is the ability of couples to have children, if they want to. The interesting fact is that while fecundability is increasing, fecundity is decreasing. Fecundability seems to be a probability that is grounded from below. It seems perfectly reasonable to speak of propensities concerning fecundability, and yet it seems unreasonable to speak of propensities concerning fecundity. If we are dealing with causal factors concerning fecundity, it appears that the propensity theory is inadequate. This is because there simply may be no propensities concerning fecundity. If we are going to use fecundity in a probabilistic causal context, we will have to have some different interpretation of probability. This will be true whenever the probabilities are not grounded from below.

However, even though it seems natural to interpret probabilistic causality in terms of probabilistic dispositions, it seems as if there are reasons for not letting these propensities actually define probability. Problems arise when we consider inverse probabilities of the sort given to us by Bayes's theorem (Salmon forthcoming, pp. 240-1). Often it seems very natural to say that event A has a certain propensity or probabilistic disposition to cause an event B, while it seems very unnatural to say that event B has a certain propensity or causal disposition to have been produced by A. Dispositions seem to have a temporal direction built into

them, which is not present in all probability statements. Thus even though propensities seem to be what we are searching for in a theory of probabilistic causality, there do seem to be problems with letting propensities be an interpretation of probability.

A.6. The Frequency Interpretation of Probability

The relative frequency interpretation of probability defines probability in terms of the frequency with which certain attributes occur in sequences of events. The probability of A given B is defined in terms of the ratio of the number of A and B's to the number of B's in the sequence. A finite frequency interpretation of probability takes this ratio to be the probability. An infinite relative frequency theory defines probability in terms of the limit of the ratios, as the size of the sequence approaches infinity. The intuition behind the relative frequency theory of probability is that the probability of an event in a sequence is a measure of how often the event occurs in that sequence. Since our notion of probability is intimately connected with how often events occur, the relative frequency interpretation of probability is one of the more fundamental interpretations of probability.

In contrast to the subjective interpretation of probability, the relative frequency interpretation is an objective theory. Thus it does not face the difficulties that the subjective interpretation faced. Unlike the logical interpretation, the relative frequency interpretation is based solely upon experience. There are no a priori factors that enter into probability relations according to the relative frequency interpretation. The relative frequency interpretation does not begin with any bias as to what are causes and what are not causes; that is determined solely from experience. This fact fits in well with Hume's ideas about causation. Thus the relative frequency interpretation does not face the problems that the logical interpretation faces.

Since the relative frequency interpretation defines probability for classes of events, the relative frequency theory handles cases in which we wish to speak of causation in terms of classes and not of individuals. In looking at general laws, such as smoking causes cancer, it appears as if we are looking at classes of events

and not singular events. In general laws it does not seem as if we are interested in single case, definite, or individual probabilities; instead it seems as if we are interested in the probabilities of classes, or indefinite probabilities. I earlier raised the possibility that some probabilities may not be grounded from below. If that is true, the relative frequency theory will be the only theory that will be able to handle such cases. If it is true that probabilities regarding fecundity in marriage are not based upon propensities, and only arise in larger populations, then the relative frequency interpretation is the natural interpretation of probability to use in such cases. Since the frequency theory is based on classes, it will be able to handle naturally cases in which causation is between classes of events.

The question now arises whether the frequency interpretation of probability can handle cases in which the probabilities are grounded from below, or in which causal relations are between individual events. Since the frequency theory only defines probability for classes, probability is not defined for individual events. In order to make the frequency theory applicable to individual events, some way of assigning probability values to individual events must be devised. This is a problem in the application of the frequency theory of probability. Given that this problem can be solved, the frequency theory should provide a good measure of causal dispositions. It seems natural that probabilistic causality is based upon probabilistic dispositions, and if this is true the frequency theory attempts to provide a good measure of these dispositions.

However, there is a serious problem that the frequency theory faces. If we accept that there are some sort of causal dispositions, then we have no guarantee that the frequencies will mirror those propensities. Skyrms (1980, pp. 29–31) presents an argument which shows that nothing is more probable than that something improbable will happen. For example, suppose we are flipping fair coins; it becomes probable that if we flip enough coins that one of them will come up all heads, no matter how many times we flip it. But since this is a fair coin, the probability of it coming up heads is $1/2$, and not 1, as the frequency theory would claim. This is a problem for the frequency theory of probability, because it defines probability in terms of these sequences. This problem must be solved in order to use the frequency interpretation in a theory of probabilistic causality.

We must conclude from the preceding discussion that in discussing

probabilistic causality we are using either a propensity or relative frequency interpretation of probability. The classical, logical, and subjective interpretations all seem to have serious philosophical difficulties when it comes to probabilistic causality.

References

1. Cartwright, Nancy. "Causal Laws and Effective Strategies." *Nous* 13 (1979), 419-437.
2. Fetzer, James and Nute, Donald. "Syntax, Semantics, and Ontology: A Probabilistic Causal Calculus." *Synthese* 40 (1979), 453-495.
3. Fetzer, James and Nute, Donald. "A Probabilistic Causal Calculus: Conflicting Conceptions." *Synthese* 44 (1980), 241-246.
4. Gasking, Douglas. "Causation and Recipes". In *New Readings in Philosophical Analysis*, Herbert Feigl, Wilfred Sellars, and Keith Lehrer, Eds., Appleton-Century-Crofts, New York, 1972, pp. 518-523.
5. Giere, Ronald N.. *Understanding Scientific Reasoning*. Holt, Rinehart, and Winston, New York, 1979.
6. Good, I. J. "A Causal Calculus I,II." *British Journal for the Philosophy of Science* 11&12 (1961), 305-17, 43-51.
7. Good, I. J. "Errata." *British Journal for the Philosophy of Science* 13 (1962), 88.
8. Good, I. J. "Some Comments on Probabilistic Causality." *Pacific Philosophical Quarterly* 61 (1980), 301-304.
9. Hacking, Ian. "Grounding Probabilities from Below." *PSA 1980* 1 (1980), 110-16.
10. Hesslow, Germund. "Discussion: Two Notes on the Probabilistic Approach to Causality." *Philosophy of Science* 43 (1976), 290-292.
11. Hume, David. *A Treatise of Human Nature, Reprinted from the Original Edition in three volumes and edited with an analytical index by L. A. Selby-Bigge*. Clarendon Press, Oxford, 1888.
12. Hume, David. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals, Reprinted from the 1777 edition with Introduction and Analytical Index by L.A. Selby-Bigge*. Clarendon Press, Oxford, 1893.
13. Humphreys, Paul W. "Cutting the Causal Chain." *Pacific Philosophical Quarterly* 61 (1980), 305-314.
14. Mackie, J. L.. *The Cement of the Universe*. Clarendon Press, Oxford, 1974.
15. Mill, John Stuart. *A System of Logic: Ratiocinative and Inductive*. Longmans, Green, and Co., London, 1916.

16. Otte, Richard E. "A Critique of Suppes' Theory of Probabilistic Causality." *Synthese* 48, 2 (1981), 167-189. Copyright © 1981 by D. Reidel Publishing Company, Dordrecht, Holland.
17. Reichenbach, Hans. *The Theory of Probability*. University of California Press, Berkeley, 1949.
18. Reichenbach, Hans. *The Direction of Time*. University of California Press, Berkeley and Los Angeles, 1956.
19. Reichenbach, Hans. "The Causal Structure of the World and the Difference Between Past and Future". In *Hans Reichenbach: Selected Writings, 1909-1953*, M. Reichenbach and R. Cohen, Eds., D.Reidel Publishing Co., Dordrecht, 1978, pp. 81-119.
20. Rogers, Ben. "Probabilistic Causality, Explanation, and Detection." *Synthese* 48, 2 (1981), 201-223.
21. Rosen, Deborah A. "In Defense of a Probabilistic Theory of Causality." *Philosophy of Science* 45 (1978), 604-613.
22. Salmon, Wesley C.. *The Foundations of Scientific Inference*. University of Pittsburgh Press, Pittsburgh, 1966.
23. Salmon, Wesley C. "Why ask 'Why'?" *Proceedings and Addresses of the American Philosophical Association* 51, 6 (1978), 683-705.
24. Salmon, Wesley C. "Probabilistic Causality." *Pacific Philosophical Quarterly* 61, 1-2 (1980), 50-74.
25. Salmon; Wesley C. "In Praise of Relevance." *Teaching Philosophy* 4, 3&4 (1981), 261-275".
26. Salmon, Wesley C. *Scientific Explanation and the Causal Structure of the World*. Forthcoming.
27. Skyrms, Brian. *Causal Necessity*. Yale University Press, New Haven, 1980.
28. Suppes, Patrick. *A Probabilistic Theory of Causality*. North-Holland, Amsterdam, 1970.
29. van Fraassen, Bas C.. *The Scientific Image*. Clarendon Press, Oxford, 1980.