Causality, Propensity, and Bayesian Networks

by Donald Gillies, University College London[*]

Abstract

This paper is concerned with the problem of the relation between causality and probability. It will try to put together two approaches to this question which arose in different contexts, and which have, hitherto, developed separately. The first of these approaches involves the *propensity* theory of probability. This was introduced by Popper and developed within the philosophy of science community. In particular there have been interesting discussions of whether propensities might be causes, and this has brought about the emergence and attempts to resolve the Humphreys' Paradox. The second approach is that of *Bayesian networks* which were introduced by Pearl and have been developed within the artificial intelligence community. Bayesian networks involve both probabilities and causes, and so are tied up with the question of how causality and probability are related. In this paper I plan to bring together propensities and Bayesian networks in the hope that this synthesis may cast some additional light on the problem of relating causality to probability.

Contents

1.Introduction

In order to explain the general plan of this paper, it is necessary to make a distinction between *determinate* and *indeterminate* causality. Let us begin with determinate causality. Suppose we have a causal law of the form A causes B, I will say that the causality involved is determinate if whenever A occurs B follows. Kant in his discussions of causality always assumes that causality is determinate. Thus he says in *The Critique of Pure Reason* (1787, B5):

' … the very concept of a cause … manifestly contains the concept of a necessity of connection with an effect and of the strict universality of the rule, ...'

1

As far as my definition of determinate causality is concerned, 'the strict universality of the rule' is all that is required, though Kant adds 'the necessity of connection with an effect'.

Hume defines cause as follows (1748, Section 60, 76):

'… 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.*'

This again is determinate causality in the sense just defined. Hume also accepts that there is a necessary connection between cause and effect, though of course his analysis of necessary connection is very different from Kant's.

We can see then that both Hume and Kant, despite their philosophical differences, accept that causes are determinate, and I think that this was indeed the standard notion of cause throughout the 18th and 19th centuries – during what we might call the Newtonian period. In the 20th century, however, there has developed a notion of indeterminate causality in which A can cause B without every instance of A being followed by an instance of B. The introduction of this new notion of causality may have been influenced by quantum mechanics, but the principal use of the new concept has been in medicine and epidemiology. Indeed the very first significant appearance of the new concept occurred in 1950 when Austin Bradford Hill made his famous, and at the time highly surprising, claim that smoking caused lung cancer. Now obviously this does not mean that everyone who smokes will contract lung cancer. Indeed less that 5% of smokers get the disease. However the probability of someone who smokes 25 or more cigarettes a day getting lung cancer is 25 times the probability of a non-smoker contracting the disease.

The introduction of indeterminate causality at once raises the question of how this notion is connected to probability – a problem which did not arise for determinate causality where certainty reigns. This problem turns out to be a very tricky one. At first it might be thought that if A causes B, then the probability of B given A would be higher than the probability of B given not-A. Plausible though this principle sounds, however, there are counter-examples to it, and it holds only under rather complicated qualifying conditions. This paper is concerned with some aspects of the relation between causality, understood in the indeterminate sense, and probability. It will try to put together two approaches to this question which arose in different contexts, and which have, hitherto, developed separately. The first of these approaches involves the *propensity* theory of probability. This was introduced by Popper and developed within the philosophy of science community. In particular there have been interesting discussions of whether propensities might be causes. The second approach is that of *Bayesian networks* which were introduced by Pearl and have been developed within the artificial intelligence community. Bayesian networks involve both probabilities and causes, and so are tied up with the question of how causality and probability are related. In this paper then I plan to bring together propensities and Bayesian networks in the hope that this synthesis may cast some additional light on the problem of relating indeterminate causality to probability. By way of introduction, I will now make some preliminary remarks about both propensities and Bayesian networks. This will then enable me to give a more detailed plan for the rest of the paper.

The propensity interpretation of probability was introduced by Popper in the late 1950's (see his 1957 and 1959). In his *Logic of Scientific Discovery* of 1934, Popper had advocated a version of the frequency theory of probability, but he grew dissatisfied with this account of probability, and felt the need to develop an interpretation of probability which differed from the frequency theory, but which still held probability to be objective. This new interpretation was the propensity theory of probability. Popper further suggested in his later book: *A World of Propensities* (1990) that propensity might be a generalisation of the notion of cause. As he put it (1990, 20): 'Causation is just a special case of propensity: the case of propensity equal to 1, …' Using the terminology introduced earlier, Popper's point seems to be that indeterminate causes are propensities, and determinate causes are propensities equal to 1. So to say, for example, that smoking causes heart disease is just to say that smokers have a propensity to contract heart disease.

This thesis of Popper's is simple and attractive, but, to evaluate it, we have to take account of the fact that at present we do not have a single propensity theory, but rather a number of significantly different propensity theories. Popper's original suggestion for a propensity theory was taken up and developed in different ways by several philosophers of science. In chapter 6 of my book *Philosophical Theories of Probability* (Gillies, 2000, 113-36), I consider three different versions of the propensity theory and try to assess their respective merits. It would be of interest to carry out the investigations of the present paper for each of these different propensity theories, but this would make the paper too long, and so I will confine myself here to my own preferred version of the propensity theory, which is developed in chapter 7 of the book (Gillies, 2000, 137-68). This will be expounded briefly in the section 2, and I will then show that, for this version of propensity theory, we cannot identify propensities with indeterminate causes in the way suggested by Popper. The difficulty here, which has become known as Humphreys' paradox, is that probabilities are symmetrical, while causes are asymmetrical.

Let us now turn to Bayesian networks. As we shall see, these can often be interpreted both in terms of causes and in terms of probabilities. At this point we can remark that the notion of indeterminate causality leads naturally to a network representation. In the determinate case, a cause leads to just one effect, which, if it is itself a cause, will lead to another single effect, and so. We thus have a straightforward linear sequence of causes and effects. Indeterminate causes, however, can have several effects. For example, smoking can cause both heart disease and lung cancer. An effect can have several indeterminate causes. Thus a heart attack could be caused by the patient smoking, eating a diet which contains a high percentage of animal fats, having a stressful life, and other factors besides. Moreover the effect of a particular cause might itself cause several further effects and so on. There is no hope of representing these relationships by a simple linear sequence, but they can be represented very naturally in a network diagram such as figure 1.

Figure 1

In this diagram let us take A = smoking, B = atherosclerosis (the hardening of the arteries), C = lung cancer, D = heart attack, E = chest pain, F = anxiety. If we interpret an arrow joining X to Y as meaning that X is an indeterminate cause of Y, then the network

of figure 1 represents a group of indeterminate causal relations which are accepted by modern medicine.

It is also easy to see how probabilities can be introduced. There is, for example, a probability of C given A, that is of a patient's getting lung cancer given that he or she smokes. I will give more precise details later on, but this indicates how both indeterminate causes and probabilities can be associated with network diagrams. Bayesian networks were introduced by Pearl, but he has always interpreted the probabilities they contain in the subjective sense, i.e. as degrees of belief. However I will argue that probabilities in Bayesian networks can also be interpreted as propensities. Moreover this interpretation enables us to carry further the investigation of how indeterminate causes are related to propensities.

In section 3 I will make this initial discussion more precise by defining the notion of Bayesian network, and comparing such networks to causal networks and probability networks. Now Bayesian networks have arrows connecting their nodes, and such arrows introduce an asymmetry which might correspond to causality. It will be remembered that Humphreys' paradox arose because probabilities are symmetrical and causes asymmetrical. However, since Bayesian networks have the built-in asymmetry generated by their use of arrows, it might seem that their use could overcome Humphreys' paradox. Unfortunately it turns out, as I will argue in section 4, that this is not the case, so that causality cannot be defined in terms of Bayesian propensity networks. Nonetheless the two concepts are connected, and I will argue in section 5 that causal networks can be used as a heuristic guide for the construction of Bayesian propensity networks.

2.A Particular Version of the Propensity Theory. Humphreys' Paradox.

The particular version of the propensity theory which will be used in this paper associates propensities with repeatable conditions. Suppose **S** is a set of repeatable conditions, and one of the outcomes of **S** is A. Then the probability of A given **S** is p, or in symbols $P(A \mid \mathbf{S}) = p$, is taken to mean that the conditions **S** have a propensity, if they are repeated a large number of times, to produce sequences in which A occurs with a frequency approximately equal to p. This type of propensity theory could be called a long run propensity theory since it deals with long sequences of repetitions. In his first version of the propensity theory (1957 and 1959), Popper did indeed associate propensities with repeatable conditions, but he also wanted these propensities to apply to the single case. This position ran into difficulties because a single case can very often be described as an instance of different sets of repeatable conditions, and these different sets will produce different values of the propensity. In my 2000 (Chapter 6, 113-36), I show that these difficulties have led to the development of different kinds of propensity theory. Some of these theories try to retain objective propensities for single events, and no longer associate propensities with repeatable conditions. In the propensity theory to be used here, however, propensities continue to be associated with repeatable conditions, but they are no longer applied to the single case. This is why it is a long run propensity theory rather than a single case propensity theory. The single case is dealt with on this approach not with objective propensities, but with subjective, degree of belief, probabilities.

A second important feature of this version of the propensity theory is that propensities are regarded as satisfying the Kolmogorov axioms, so that we can speak of a propensity interpretation of the standard mathematical theory of probability. It follows from this that propensities cannot be identified with indeterminate causes as Popper suggested. The problem here was first noticed by Humphreys, and first published, with a reference to Humphreys, by Salmon in his 1979. It was named 'Humphreys' paradox' by Fetzer in his 1981, and has given rise to much interesting discussion, including an important article by Humphreys himself (1985). The argument can be put as follows. Causes are asymmetrical in the sense that if A causes B, then, in general, B does not cause A. Indeed if A causes B, then A usually occurs before B. However, apart from a few speculations in theoretical physics, it is universally conceded that causes do not operate backwards in time. So if A occurs before B, B cannot cause A. Moreover the asymmetry of causation continues to apply even for simultaneous causation. For example although a tumour causes pain, it is not the case that the pain causes the tumour. Although smoking causes heart disease, it is not true that heart disease causes smoking, and so on. The situation is very different with probabilities. In general if $P(A \mid B)$ is defined, then so is $P(B \mid A)$. Suppose then that we give the probabilities a propensity interpretation and regard $P(A \mid B)$ as stating that B causes A in the indeterminate sense. It unfortunately follows that if B causes A in this sense, then also A causes B. This contradicts the asymmetry of causation, and shows that we cannot identify propensities with indeterminate causes.[1]

It should be noted that this argument makes essential use of the assumption that propensities obey the Kolmogorov axioms. Fetzer who wants to identify propensities with indeterminate causes has suggested that propensities might obey a different set of axioms. In fact with Nute he has developed a probabilistic causal calculus (see his 1981, 59-67). Another way out of the difficulty might be to make the propensity interpretation only partial so that some probabilities are propensities, while others are not. If however propensity is regarded as a complete interpretation of the standard mathematical calculus, as is done here, then propensities cannot be identified with causes. Yet even if propensities are not always causes, it seems plausible that there should be some link between causes and propensities. In the next section I will begin an examination of the nature of this link in the context of networks.

3. Causal Networks, Probability Networks, and Bayesian Networks

The terms 'causal network' and 'Bayesian network' are often used interchangeably. Since the aim of this paper is to examine the relations between causality and probability, more precision is needed, and I will begin by defining the concepts of causal network, probability network and Bayesian network in a precise fashion which will show that they are different. We shall need further to distinguish between Bayesian belief networks and Bayesian propensity networks. After these complicated, but necessary distinctions which will be made in the present section, I will proceed in the last two sections of the paper to examine the relationship between causality and propensity in

the present context. This means, in effect, examining how causal networks are related to Bayesian propensity networks.

Let us begin by defining the concept of network. This is done using the concepts of graph theory. A *directed graph* can be defined as a finite set of nodes, some pairs of which are joined by an arrow. We will use capital letters, A,B, C, X, Y, Z, … to stand for nodes. In the diagrams these will be surrounded by circles. A *cycle* in a directed graph is a sequence of nodes $C_1, C_2, \ldots, C_n$, such that there is an arrow between C_1 and C_2, C_2 and C_3, …, C_{n-1} and C_n, and C_n and C_1. If a directed graph contains no cycles, it is said to be *acyclic*. A *network* is defined to be a directed acyclic graph (or DAG). Figure 1, which we considered earlier, gives a diagram of a network.

If an arrow runs from A to B in a network, A is said to be a *parent* of B, and B a *child* of A. Parents, parents of parents, and so on are said to be *ancestors*; while children, children of children, and so on are said to be *descendants*. So in figure 1, A has children B, C and D, and descendants B, C, D, E and F. E has parents B and C, and ancestors A, B and C.

A *causal network* can now be simply defined as a network in which an arrow joining A to B means that A causes B. It will be seen at once that this is a useful concept for dealing with indeterminate causality since, with this notion of causality, many effects have several causes, and many causes several effects. Moreover the use of arrows is a very good way to represent the asymmetry of causality.

Suppose the nodes of a network are random variables having a joint distribution. Then the network will be called a *probability network*. A *Bayesian network* is a probability network which satisfies the following *independence conditions*. These state that a node is conditionally independent, given its parents, of any other set of nodes in the network not containing any of its descendants. We could say that in a Bayesian network, the parents of a node *screen it off* from the other nodes of the network except its descendants. Now the use of the term 'screening off' will naturally suggest to philosophers some ideas of Reichenbach's which have been developed by Salmon. In his discussion of causality in 1956, Reichenbach introduced the notion of conjunctive fork, and spoke of the common cause in such a fork as screening off one of its effects from the other. His ideas are developed in Salmon (1978). In fact Reichenbach's definition of conjunctive fork is just a special case of the independence conditions which define a Bayesian network. Since Reichenbach was concerned with the analysis of causality, this strongly suggests that there must be a link between Bayesian networks and causality. There is indeed such a link, but the nature of the link is not at all straightforward as I will now go on to show.

From the above definitions, it is obvious that causal networks and Bayesians networks are different. A causal network is a set of causes and effects arranged in a network. One can specify a causal network without mentioning probabilities or random variables. A Bayesian network is a set of random variables with a joint distribution arranged in a network and satisfying the independence conditions. One can specify a Bayesian network without mentioning causes.

6

Despite the difference between the two types of network, it is possible to go from one type to another. Let us consider the case of going from a causal network to a probability network. Start with a causal network. Suppose it is possible to interpret the nodes as random variables on the same probability space, and that we can assign the variables a joint distribution. We then have *a* probability network *corresponding* to the causal network. Note that here we have *a* rather than *the* corresponding network, because, starting with a causal network, it might be possible to assign different probability distributions to produce different probability networks. This is connected to the question of the interpretation of probability in Bayesian networks to which I now turn.

The theory of Bayesian networks was developed in the 1980's by Pearl (1986 and 1988) and Lauritzen and Spiegelhalter (1988). Pearl has always interpreted the probabilities in the networks subjectively as degrees of belief. Thus he says (1988, 29):

'Bayesian methods provide a formalism for reasoning about partial beliefs under conditions of uncertainty. In this formalism, propositions are given numerical parameters signifying the degree of belief accorded them under some body of knowledge, and the parameters are combined and manipulated according to the rules of probability theory.'

For this reason, Pearl uses the expression '*Bayesian belief networks*' (1988, 50). Pearl continues to hold the same point of view in his most recent book, writing (2000, 2):

'We will adhere to the Bayesian interpretation of probability, according to which probabilities encode degrees of belief about events in the world and data are used to strengthen, update, or weaken those degrees of belief.'

Lauritzen and Spiegelhalter were also working in the tradition of subjective Bayesianism, but they seem less definitely committed to this view than Pearl. This is what they say (1988, 159):

'Our interpretation of probabilities is that of a subjectivist Bayesian …. This seems a convenient and appropriate view in an area concerned with the rational structuring and manipulation of opinion, and the subjectivist objectives of a coherent system of probabilities representing belief in verifiable propositions, successively updated on the basis of available evidence, appears to fit remarkably the objectives of expert systems research. However, many of the techniques presented here are appropriate in disciplines where graphical structures are used and a frequentist interpretation is more appropriate, such as complex pedigree analysis in genetics.'

So Lauritzen and Spiegelhalter think that, in some cases at least, the probabilities in Bayesian networks might be given an objective interpretation. Neapolitan (1990) is also favourable to objective probabilities in Bayesian networks. In Sucar, Gillies, and Gillies, 1993, a Bayesian network is used to resolve a practical problem in endoscopy. The probabilities in the network are interpreted as propensities in the sense explained in this paper. Now it might be thought that it makes little difference from a practical point of view whether the probabilities in a Bayesian network are interpreted subjectively or objectively, but this is not so. The use of subjective probabilities, as Pearl indicates in the passages just quoted, is connected with a Bayesian methodology in which evidence is used to update degrees of belief through a process of Bayesian conditionalisation. The

use of objective probabilities is similarly connected with a Popper (or Popper-Fisher) methodology similar to that of classical statistics. Here the emphasis is on the testing of any assumptions made by means of statistical tests. As applied to Bayesian networks, this means that tests should be carried out to see whether the independence conditions are satisfied, and that the network must be modified if the independence conditions do not hold. In Sucar, Gillies, and Gillies, 1993, it is shown that this testing methodology can result in simplified Bayesian networks which perform better. A discussion of Bayesian methodology in the context of Bayesian networks is to be found in my 1998. I will come back to these issues in section 6, but will not consider the issue of Bayesianism further here, beyond remarking that the term 'Bayesian network' as used in this paper does not imply either the interpretation of probabilities as degrees of belief, or the use of a Bayesian methodology. Of course it does contain an appropriate reference to Bayes' theorem, but it should be remembered that this is a theorem of the probability calculus, and holds whether that calculus is interpreted subjectively or objectively.

We are now in a position to distinguish between two types of probability networks. These are (i) *belief networks*, in which the probabilities are interpreted subjectively as degrees of belief, and *(ii) propensity networks*, in which the probabilities are interpreted objectively as propensities. Corresponding to a causal network, there might be several different belief networks specifying the different degrees of belief of different individuals. If, however, we are considering propensity networks, the situation is rather different. In order to move from a causal network to a propensity network, we have to specify an underlying set of repeatable conditions **S**. We must further interpret the causes in the causal network, not as referring to particular instances, but as general causal laws holding for different repetitions of the underlying conditions **S**. Suppose the set of elementary or atomic outcomes of **S** is □, and that we can interpret the nodes of the causal graph as random variables defined on □. There is then just one propensity network corresponding to the causal network, since propensities are objective and have their values fixed by the nature of the external world. In general, assuming tacitly as given an underlying set **S** of repeatable conditions relative to which a propensity interpretation is possible, we shall speak of *the* propensity network corresponding to a given causal network, since at most one such network can exist in these circumstances.

We can now define a *Bayesian belief network* as a belief network in which the independence conditions are satisfied. This agrees with Pearl's use of the expression which we quoted earlier. We can further define a *Bayesian propensity network* as a propensity network in which the independence conditions are satisfied. This concludes our sequence of preparatory definitions, and we are now in a position to return to the question of how causality is related to propensity and to consider it in the context of Bayesian networks. This investigation will occupy the last two sections of the paper.

4.Can the use of Networks overcome Humphreys' Paradox?

In section 2 we considered the possibility that propensities could be identified with causes, and indeed could be considered as a kind of explication of causes. This suggestion did not work, however, because of the problem that propensities are symmetrical, and causes asymmetrical - the difficulty known as Humphreys' paradox. We can now raise a related question by asking whether Bayesian propensity networks

might be identified with causal networks, and could be considered as a kind of explication of causal networks. Now Bayesian propensity networks do involve arrows and hence are not symmetrical. Thus it might seem plausible that they could overcome the symmetry/asymmetry difficulties of Humphreys' paradox. Yet, in this treacherous field, plausible suggestions often turn out to be false, and this holds again in the present case.

The problem is that it is fairly easy to construct examples of Bayesian propensity networks, for which the corresponding network is not a causal network. To see this let us start with the very simplest case (see figure 2(a)) of a Bayesian propensity network consisting of two nodes A and B joined by an arrow from A to B.

Figure 2

Let us suppose that A causes B. Then the Bayesian propensity network of figure 2(a) is also a causal network. But now consider the network obtained by making the arrow run from B to A rather than from A to B (see figure 2(b)). Suppose we assign the same joint propensity distribution to the variables A, B as in the previous case. The new network is still a Bayesian propensity network, since the independence conditions are automatically satisfied. However, in this case, it no longer corresponds to a causal network, assuming, as is characteristically the case, that, although A causes B, B does not cause A, i.e. this is an asymmetrical case of causation.[2]

A little reflection shows that the underlying problem is still that of Humphreys' paradox, although it appears in a somewhat different form in the case of networks. There is indeed an asymmetry in a Bayesian propensity network introduced by the direction of the arrows. However, it is always possible to reverse the direction of one or more arrows in such a propensity network, and then by making suitable adjustments to obtain another Bayesian propensity network. The situation is very different with causal networks. Suppose, as is usually the case, that all the causes involved in a particular causal network are asymmetrical. If we then reverse the direction of any of the arrows of the network, it will cease to be a causal network, and we will not be able to turn it back into a causal network by adjustments elsewhere in the network. So far I have illustrated these points by the very simple example of a two node network, which might indeed be a special case. I will now give an example of a three node network in which the same phenomena are to be found, and it will be obvious that this can be generalised to any number of nodes.

The example of a three node network is given by Pearl, 1988, 116, and is discussed by Neapolitan, 1990, 187-9. Consider the following set of repeatable conditions. Toss independently two fair coins whose results are given by A and B. If either of them gives heads, ring a bell. The variable C refers to the bell. C = 1 if the bell is rung, and C = 0 if it is not rung. We can represent this simple repeatable experiment by the network shown in figure 3(a).

Figure 3

The network of figure 3(a) is a Bayesian propensity network relative to the repeatable conditions defining the experiment. We can easily write down the propensity distributions of A, of B, and of C given A and given B. Here the independence

9

conditions only require that A and B are independent which is the case. The network of figure 3(a) is obviously also a causal network, since the results of the coin tosses determine whether the bell is rung or not. Let us now alter the network by reversing the arrows joining A to C, and B to C, while also adding an arrow joining A to B. The result is shown in figure 3(b). Let us assign to this network the same joint propensity distribution as before. We still have a Bayesian propensity distribution. If we had not joined A to B with an arrow, then the propensity network would no longer have been a Bayesian propensity network, since the independence conditions in this case would have required that A should be independent of B conditional on C, which would not have been the case. If we add an arrow joining these two conditionally dependent nodes, however, we ensure that the independence conditions are satisfied in the resulting network, which is a Bayesian propensity network. The network of figure 3(b) is obviously, however, no longer a causal network. The ringing or silence of the bell has no causal influence on whether the first coin earlier came down heads or tails.

It is easy to see how this example can be generalised. Take any Bayesian propensity network which corresponds to a causal network. Then reverse one or more arrows, while retaining the same joint propensity distribution. The result will be a propensity network which may no longer be a Bayesian propensity network because some of the independence conditions may no longer be satisfied. However we can restore the independence conditions by adding arrows between nodes which are conditionally dependent when they should be conditionally independent for the independence conditions to be satisfied. In this way we generate a Bayesian propensity network with the same nodes and joint propensity distribution as the previous one, but which is no longer a causal network.   This shows beyond doubt that we cannot identify causal networks with Bayesian propensity networks. Once again, the concepts of causality and propensity show themselves to be distinct in some fundamental respects.

5.Causal Networks as a Heuristic Guide for the Construction of Bayesian Propensity Networks

In section 4 it was shown that if a network is a Bayesian propensity network, it does not necessarily follow that it is a causal network. But what about the converse? Suppose then that we have a causal network for which there is a corresponding propensity network. Can we infer that this propensity network is a Bayesian propensity network, i.e. that the independence conditions are satisfied? I will show in this section that the answer is: 'no', by presenting three counter-examples to the thesis that a propensity network corresponding to a causal network is a Bayesian propensity network.

The first counter-example could be called the endoscopy example for it arose in the attempt to construct an expert system for colon endoscopy. It is described in Sucar, Gillies, and Gillies, 1993. In colon endoscopy, a camera, which transmitted pictures to a television screen, was inserted into the patient's colon, and steered along the colon. In moving the endoscope, it was very important to direct its tip towards the opening of the next section of the colon which is called the *lumen*. If this was not done correctly it could be very painful and dangerous for the patient, and could even cause perforation of the colon wall. One problem then was to pick out from the picture being transmitted by the endoscope's camera the position of the lumen. A domain expert used his or her

judgement, based on experience, to decide where the lumen was located. The aim of the expert system was to automate this process. Consultations were carried out with an expert endoscopist to find out what features were used to pick out the lumen. It turned out that in different situations different features indicated the lumen. For example several concentric rings could indicate the lumen, or a large, uniform dark region could indicate the lumen. On the basis of this information a causal network was constructed. Figure 4 gives a part of this network.

Figure 4

L stands for the lumen, which causes a large dark region (or LDR) to appear on the screen. This in turn produces values for the variables S, which measures the size of the region in pixels, M which measure its mean intensity, and V which measures the variance of that intensity. There was an abundant supply of videotapes of colon endoscopies in each frame of which the lumen could be indicated by the expert. These gave a mass of frequency data from which the propensities in the corresponding propensity distribution could be estimated. It is worth noting that no use was made of subjective degree of belief probabilities in this process. Having found the propensity network corresponding to figure 4 in this way, the question could now be raised as to whether it was a Bayesian propensity network, i.e. whether the independence conditions were satisfied. These conditions implied that M and V should be independent given LDR. However empirical tests revealed that this was not the case, and that M and V were in fact strongly correlated given LDR. In this case, therefore, the propensity network corresponding to a causal network was not a Bayesian propensity network.

The explanation of the correlation of M and V was discovered by Williamson (see his 2001, Section 2.2, 84), and turned out to be very simple. There is in fact a well-known mathematical relation between M and V. Let X be a random variable, and E signify expectation. Then M (the mean of X) = $E(X)$, and V (the variance of X) = $E(X^2) - [E(X)]^2 = E(X^2) - M^2$. There is thus a mathematical relation between M and V, and this accounts for the observed correlation. Generalising we can say that if mathematical relations hold between some of the nodes of a causal network, these relations could produce correlations which prevent the corresponding propensity network being a Bayesian propensity network, i.e. the independence conditions holding. Since there can often be mathematical relations of a subtle and non-obvious form, this demonstrates the need for testing to see whether any assumed independence relations do actually hold.

The second counter-example could be called the pregnant cow example, and is given in Jensen, 1996, 36-7. It comes from the field of veterinary science, and is concerned with testing to see if a cow is pregnant. (One of the charms of studying Bayesian networks consists in the curious range of things to which they apply!) A causal network corresponding to the test is given in figure 5(a).

Figure 5

Here Pr is a variable which is 1 if the cow is pregnant and 0 otherwise. BT stand for the result of a blood test, UT for the result of a urine test, and Sc for the result of scanning. Once again there is ample data to estimate propensities, and produce a corresponding propensity network. The independence conditions here state that the variables BT, UT,

and Sc should be conditionally independent given Pr. Now Sc is indeed conditionally independent of BT and UT, given Pr, but BT and UT are correlated given Pr. This is a Reichenbachian case, because the correlation between BT and UT can be explained by a common cause. Both tests depend on the hormonal level of the cow, which obviously does not affect the result of scanning. If we add an extra node Ho as shown in figure 5(b), then the propensity network corresponding to this new causal network is a Bayesian propensity network.

This example gives a general reason why a causal network might not give rise to a propensity network which satisfies the independence conditions. Any such network produced by a human to model some situation in the world could very well omit some important cause. If this omitted cause is a common cause explaining the correlation between two nodes in the network, then the corresponding propensity network could well fail to satisfy the independence conditions. It might at first be thought that one could avoid this difficulty by insisting that the causal network be complete, and not leave out any causes. However the notion of the completeness of a causal network is one which really does not make much sense. Consider any causal law of the form A causes B. We can nearly always find a causal chain $C_1, C_2, \ldots C_n$ connecting A to B, i.e. such that A causes C_1, C_1 causes C_2, $\ldots$, C_n causes B. One consequence of this is that, if A causes B, we can nearly always find some intermediate cause C, such that A causes C and C causes B. The introduction of such intermediate causes can go on virtually indefinitely. Indeed Salmon (1980, 224) quotes with approval John Venn's statement that it might be better to use the expression 'rope of causation' rather than 'chain of causation'. All this means, however, that, given any causal network, we can always extend it by adding intermediate nodes, so that the concept of a complete causal network does not seem to be appropriate. Of course in many cases the addition of intermediate nodes might be pointless, while in other cases, such as our pregnant cow example, the addition of an intermediate node is essential in order to ensure that the independence conditions are satisfied. In a particular case, there seems be no *a priori* method for telling whether additional intermediate nodes are needed or not. Consequently we should when generating Bayesian networks from causal networks always test out the independence assumptions to see whether they hold, and then adjust the network appropriately if these assumptions are not satisfied.

The third counter-example could be called the Compton scattering example. It comes from Salmon, 1978, 133-4. The context is one in which Salmon is developing critically Reichenbach's notion of a *conjunctive fork*. Suppose we have two nodes A and B which are correlated, and this correlation is explained by the fact that they are both effects of a common cause C. Suppose further that C gives a complete explanation of the dependency between A and B in the sense that A and B are probabilistically independent give C. Then the causal network with arrows joining C to A and C to B (see figure 6) is called by Reichenbach a conjunctive fork. Now a conjunctive fork in Reichenbach's sense is a Bayesian network, although the notion of Bayesian network had not been introduced when Reichenbach was writing. Indeed Reichenbach's notion of a conjunctive fork can be considered as one of the precursors of the concept of Bayesian network.

Figure 6

Salmon, in his further studies of causality, found it necessary to introduce, in addition to Reichenbach's conjunctive fork, a second kind of causal fork, which he called an *interactive fork*. As he says (1978, 134):

'It thus appears that there are two kinds of causal forks: (1) Reichenbach's *conjunctive fork*, in which the common cause screens off the one effect from the other, …, and (2) *interactive forks*, exemplified by the Compton scattering of a photon and an electron.'

The example of Compton scattering is indeed an interesting one, and I will expound it using figure 7 which is taken from Salmon, 1978, 133.

Figure 7

In a Compton scattering experiment, an energetic photon collides with an electron which can be regarded as more or less stationary. The collision is represented by the node C, where the variable C has the energy E as its value. As the result of the collision we get an electron with energy E_1 represented by node A, and a photon with energy E_2 represented by node B. Now because of the conservation of energy, we have $E = E_1 + E_2$, and so A and B are highly correlated given C. C is a common cause of A and B, but it does not screen off A from B. So this is a causal fork which is not a conjunctive fork. It is what Salmon calls an interactive fork.

Compton scattering is a quantum phenomenon and inherently probabilistic. If we repeated the experiment with the same value for E, we would get different values for E_1 each time. We can thus associate with the causal network of figure 7, a propensity network, but this propensity network does not satisfy the independence conditions, and so is not a Bayesian network. Moreover it does not look as if we could convert the network of figure 7 into a Bayesian propensity network by introducing an intermediate cause, as an additional node, as was done in the pregnant cow example. We can, however, convert the network of figure 7 into a Bayesian propensity network by adding an arrow from A to B as was done in figure 3(b). This Bayesian propensity network, however, no longer corresponds to a causal network, since there is no causal link between A and B.

In the light of these counter-examples, it is clear that the propensity network corresponding to a causal network will not in general be a Bayesian propensity network. Moreover there do not appear to be any plausible restrictive conditions on a causal network which ensure that the corresponding propensity network is Bayesian. When discussing the pregnant cow example, we pointed out that it is nearly always possible to interpolate a further cause between a given cause and its effect. It is this which makes it impossible to formulate any kind of completeness condition for causal networks, but such a completeness condition would seem to be necessary to ensure that the corresponding propensity network is Bayesian.

In spite of all these problems there is still an important link between causal networks and Bayesian propensity networks. Causal networks can be very useful heuristic guides for the construction of Bayesian propensity networks in the context of the *qualquant* methodology. This methodology for constructing expert systems was introduced in Sucar, Gillies, and Gillies (1993), and I will now explain its main features.

Qualquant is short of *qual*itative orientation leading to *quant*itative improvement. It is based on three principles. The first of these principles is that as far as possible only qualitative suggestions should be sought from the domain expert, and it should be left to the computer scientist to give this a more precise quantitative form. In practice this usually means that the constructor of the expert system should obtain a causal network from the expert. Experts in the medical field are very familiar with the cause and effect relations which hold in a particular area, and can easily transmit these to the expert system builder in the form of a causal network. Experts are not so good, however, at quantifying personal, subjective probabilities, and this gives rise to the second principle of the qualquant methodology, namely that objective probabilities should be used wherever possible. The idea in short is to elicit the basic causes from the domain experts, but to estimate probabilities from data. This naturally restricts the qualquant methodology to situations in which abundant data is available, but there are many such cases as our endoscopy and pregnant cow examples showed. The third and perhaps most important principle is the testing principle, namely that all assumptions should be tested, and modified if they fail the test. The need for this is clearly shown by the counter-examples of the present section. When we obtain a propensity network from the causal network supplied by the expert, we cannot be sure that it will be a Bayesian propensity network, i.e. that the independence conditions hold. Statistical tests must be carried out to see whether these independence hold, and the network modified if they do not hold. The use of the objective probabilities and the testing principle means that this methodology is in accordance with Popper's theory of scientific method, and also with the classical approach to statistics of Fisher and Neyman-Pearson. The qualquant methodology differs however from the subjective Bayesian approach, which is characterised by the use of subjective or degree of belief probabilities, and the absence of testing.

Suppose we are applying the qualquant methodology. A causal network has been obtained from the domain expert. The corresponding propensity network has been calculated, but testing has shown that some of the independence conditions are not satisfied. The question now arises as to what strategies should be applied to modify the propensity network in order to turn it into a genuine Bayesian network. In fact in the development of the qualquant methodology, three strategies have emerged, which correspond loosely with the three counter-examples given in this section.

The first which was used in the original paper (Sucar, Gillies, and Gillies, 1993), and is the simplest. Suppose statistical tests have revealed that two nodes are conditionally correlated which should be conditionally independent if the network is to be a Bayesian network. The strategy is to eliminate one of the nodes on the grounds that, just because the two nodes are correlated, we can get almost as much information from just one as from the pair. In the endoscopy example, the mean (M) and variance (V) were found to be correlated as a result of testing. It turned out that eliminating either of them from the network, produced a network which was not only simpler but gave improved performance.[3]

The second strategy is Reichenbachian in character. Suppose two nodes (A, B say) are conditionally correlated which should be conditionally independent if the network is to be a Bayesian network. We then try to introduce a new node C which is a

common cause of the two nodes such that the causal network illustrated in figure 6 is a conjunctive fork. This strategy is exemplified by the introduction of a node for hormone level in the pregnant cow example. A general method for introducing such hidden nodes is given in Kwoh and Gillies, 1996.

The third strategy is known as 'adding arrows' and was introduced by Wiliamson.[4] Suppose again that two nodes A and B are conditionally correlated which should be conditionally independent if the network is to be a Bayesian network. We then simply join them with an arrow. We described earlier how this could be done in the case of the interactive fork produced by Compton scattering.

There is, however, one important respect in which the technique of adding arrows differs from the previous two techniques. If a Bayesian propensity network is generated using the first two techniques, then this network will still be a causal network. If we delete a node, or add a hidden node which is a common cause, we are not disturbing the causal character of the network. Adding arrows is different, however, because the added arrow may not correspond to a causal link, as indeed was the case in the Compton scattering example. By using the strategy of adding arrows, therefore, we may often generate Bayesian propensity networks which are no longer causal networks. This might be used as an argument for not employing the adding arrows technique, but such a conclusion seems to me unjustified for reasons which I will now explain.

The point in favour of the adding arrows technique is that it gives a very simple method for generating Bayesian propensity networks. The advantage in turn of Bayesian networks is that they allow a convenient set of mathematical and computational techniques to be employed in particular instances for calculating probabilities. Moreover if we use adding arrows starting from a causal network supplied by a domain expert, then in the resulting Bayesian propensity network, the arrows corresponding to the original causal network still have a causal interpretation. Thus at least some, even if not all, of the generated network retains the intuitively satisfactory causal interpretation, while the network now has a form which enables computations to be carried out easily. Thus a judicious use of the adding arrows technique can result in a good balance between the requirements of being humanly comprehensible and of being convenient for the computer. There is always a tension between these two requirements, so that a reasonable balance between them is what we should seek.

Notes

1.This is rather a terse statement of the argument. It is developed in greater detail with examples in my 2000, Chapter 6, 129-36.

2.This example is discussed from a related, though somewhat different, point of view in Neapolitan, 1990, 186.

3.For details, see Sucar, Gillies, and Gillies, 1993, Table 3, 206.

4.An exposition of the adding arrows approach is given in Williamson, 2001, Sections 4, 5 and 6, 96-112. This, however, is in the context of Williamson's two stage methodology which differs from the qualquant methodology given here in being more subjective in character and also using maximum entropy and minimum cross entropy techniques.

References

Fetzer, J.H. (1981). *Scientific Knowledge: Causation, Explanation, and Corroboration*, Boston Studies in the Philosophy of Science, Vol. 69, D.Reidel, Dordrecht.

Gillies, D.A. (1998). Debates on Bayesianism and the Theory of Bayesian Networks, *Theoria*, **64**(1), 1-22.

Gillies, D.A. (2000). *Philosophical Theories of Probability*. Routledge.

Hume, D. (1748). *Enquiry concerning the Human Understanding*. Selby-Bigge Edition, Oxford, 1963.

Humphreys, P. (1985). Why Propensities cannot be Probabilities, *The Philosophical Review*, **94**, 557-570.

Jensen, F.V. (1996). *An Introduction to Bayesian Networks*. UCL Press.

Kant, I. (1787). *Critique of Pure Reason*. Second Edition. English translation by Norman Kemp Smith, Macmillan, 1958.

Kwoh, C.-K. and Gillies, D.F. (1996). Using Hidden Nodes in Bayesian Networks, *Articificial Intelligence*, **88**, 1-38.

Lauritzen, S.L. and Spiegelhalter, D.J. (1988). Local Computations with Probabilities on Graphical Structures and their Application to Expert Systems (with discussion), *Journal of the Royal Statistical Society B*, **50**(2), 157-224.

Neapolitan, R.E. (1990). *Probabilistic Reasoning in Expert Systems*. John Wiley.

Pearl, J. (1986). Fusion, Propagation and Structuring in Belief Networks, *Artificial Intelligence*, **29**, 241-288.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.

Pearl, J. (2000). *Causality. Models, Reasoning, and Inference*. Cambridge University Press.

Popper, K.R. (1934). *The Logic of Scientific Discovery*. 6th revised impression of the 1959 English translation. Hutchinson. 1972.

Popper, K.R. (1957). The Propensity Interpretation of the Calculus of Probability, and the Quantum Theory. In S. Körner (ed.) *Observation and Interpretation, Proceedings of the Ninth Symposium of the Colston Research Society, University of Bristol,* Butterworths Scientific Publications, 65-70 & 88-9.

Popper, K.R. (1959). The Propensity Interpretation of Probability, *British Journal for the Philosophy of Science*, **10**, 25-42.

Popper, K.R. (1990). *A World of Propensities*. Thoemmes.

Reichenbach, H. (1956). *The Direction of Time*. University of California Press.

Salmon, W.C. (1978). Why Ask, "Why?"? An Inquiry Concerning Scientific Explanation, in Salmon, 1998, Ch. 8, 125-41.

Salmon, W.C. (1979). Propensities: a Discussion Review, *Erkenntnis*, **14**, 183-216.

Salmon, W.C. (1980). Probabilistic Causality, in Salmon, 1998, Ch. 14, 208-232.

Salmon, W.C., (1998). *Causality and Explanation*, Oxford University Press.

Sucar, L.E., Gillies, D.F., and Gillies, D.A. (1993). Objective Probabilities in Expert Systems, *Artificial Intelligence*, **61**, 187-203.

Williamson, J. (2001). Foundations for Bayesian networks. In D. Corfield and J. Williamson (eds.), *Foundations of Bayesianism*, Kluwer, 75-115.