Philosophical Foundations for Causal Networks

Glenn Shafer

Department of Accounting and Information Systems
Faculty of Management, Rutgers University
180 University Avenue, Newark, New Jersey 07102 USA
gshafer@andromeda.rutgers.edu.

Abstract. Bayes nets are seeing increasing use in expert systems [2, 6], and structural equations models continue to be popular in many branches of the social sciences [1]. Both types of models involve directed acyclic graphs with variables as nodes, and in both cases there is much mysterious talk about causal interpretation. This paper uses probability trees to give precise conditions under which Bayes nets can be said to have a causal interpretation. Proofs and elaborations are provided by the author in [4].

1 Introduction

In spite of the impression of action conveyed by its arrows, a Bayes net is not necessarily a dynamic model. It merely expresses conditional independence relations for a single probability distribution, and this is static information. So what do we mean when we claim causal status for a Bayes net?

When we say that a Bayes net is causal, we are apt to elaborate by saying that its arrows point in the direction of time and causality. When there is an arrow from variable X to variable Y, X precedes and causes Y. But this remains a bit mysterious. Perhaps it means that (1) there is a point in time where X is determined and Y is not yet determined, and (2) if the value of X at that point is x, then the probabilities for Y at that point are the conditional probabilities for Y given X=x. But functional dependence, which is generally accepted as an important example of causality and is often used in Bayes nets, violates condition (1). If Y is a function of X, say $Y=X^2$, then Y is determined at least as soon as X and perhaps sooner.

This article uses an overtly dynamic framework for probability—the probability tree—to clarify these issues. A Bayes net can be thought of as an incomplete description of a probability tree, and we may say that a Bayes net is causal if it is a incomplete description of nature's probability tree.

As it turns out, there is more than one way a Bayes net can usefully describe nature's probability tree. We may choose to draw a net with the property that whenever there is an arrow from X to Y, X is always determined at least as soon as Y as nature moves through the tree. Or we may instead draw a net such that the arrows to Y indicate that Y is a function of the variables from which the arrows come, in which case these variables—some or even all of them—may be determined after Y.

In the probability-tree framework, temporal precedence and functional dependence are unified by the concept of *tracking*. We say that X tracks Y if the probabilities for Y are the same at any two points in the tree where X is just settled to have a particular value, say x. This definition allows for the possibility that Y itself is already settled at all the points where X has just been settled, in which case the probabilities at those points are zeros and ones, and Y is indeed a function of X. It also allows for the possibility that Y is not yet settled at any of these points, in which case X may be said to precede Y. It even allows for the possibility that X may be settled sometimes before Y and sometimes after Y.

The word "cause" can be used in several ways in a probability tree, but it when we are speaking of nature's tree, it is most natural to say that steps in the tree are causes of how things come out in the end. The causes of a variable Y are the steps that change Y's probability. When we use "cause" in this way, we may say that tracking is a causal concept. As it turns out, X tracking Y means that until X is settled Y has no causes other than the causes of X.

Some readers will want to say "X causes Y" instead of "X tracks Y," but this is a recipe for confusion and even mischief. Truth be told, variables are not causes. Things that happen and things one does are causes. The total number of cigarettes smoked is not a cause of lung cancer. It is the smoking—the steps in nature's tree where one smokes or where things happen that lead to more smoking—that raise the probability of lung cancer and hence can be said to be among its causes. The total number of cigarettes measures the effect of these steps, but this might also be measured by the stain on one's fingers. Moreover, tracking is only one of many concepts that we can use to relate causes of one variable to causes of another.

Another very important causal concept is *independence*. Two variables X and Y are independent in the probability tree sense if they have no common causes—i.e., if there are no steps in the tree that change both the probabilities for X and the probabilities for Y. This definition of independence is stronger than the usual one; if X and Y are independent in this probability-tree sense, then their probabilities multiply at every point in the tree.

The thesis of this article is that one way to interpret a Bayes net causally is to assert two things about nature's probability tree: (1) each variable is tracked by its parents, and (2) after its parents are settled, it is independent of its nondescendants.

The ideas outlined here are explained more fully the author's book, *The Art of Causal Conjecture* [4] which also develops the probability-tree understanding of causality in several other directions. It uses concepts analogous to tracking (tracking in mean, linear tracking, and linear sign) and generalizations of independence (uncorrelatedness and unpredictability in mean) to account for structural equation models and path diagrams, probability structures analogous to Bayes nets that have been used widely in the social sciences. It also shows how the ideas of tracking, sign, and uncorrelatedness generalize from probability trees and Bayes nets to decision trees and influence diagrams.

2   Situations, Events, and Variables in a Probability Tree

A probability tree represents a plan for performing a sequence of experiments. The plan specifies the possible results of each experiment. At each step, it also specifies what experiment is to be performed next, depending, perhaps, on how earlier experiments have come out. As we perform the sequence of experiments, we follow some path through the tree. Figure 1 shows an example. Here we begin by spinning a fair coin. If it comes out heads, we spin another fair coin. If it comes out tails, we spin a coin biased to heads, and so on.

The sequence of experiments can be thought of as a composite experiment. The stop signs constitute the sample space of this composite experiment. Any subset of the sample space is an event. The circles and stop signs are *situations*. The circles are situations in which an experiment is performed, and the stop signs are terminal situations—situations in which experimentation has stopped. The initial situation is designated by Ω.

Every situation corresponds to an event—the set of stop signs below it. In Figure 2, for example, the situation W corresponds to the event {a,b,c}.) We often find it convenient to identify the situation with this event. Most events, however, are not situations.

When an event E contains a situation S, we say that E is *certain* in S. In Figure 2, {a,d,e} is certain in U. When E∩S=∅, we say that E is *impossible* in S. In Figure 2,

{k,m} is impossible in W. When E∩F∩S=∅, we say that E and F are *incompatible* in S. When E is impossible or certain in S, it is *determinate* in S. Otherwise, it is *indeterminate* in S.
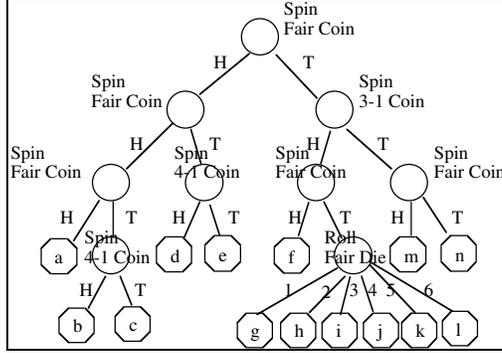
Fig. 1. A probability tree.
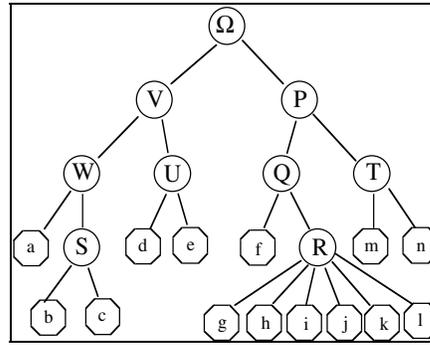
Fig. 2. Names for the situations.

Except for the initial situation Ω, every situation has a situation immediately above it, which we call its *mother*. We say that E *happens at* S if E is certain in S but not in S's mother. We say that E *fails at* S if E is impossible in S but not in S's mother. In Figure 2, the event {a,d,e} happens at {a} and at U and fails at S and at P. We say that S *resolves* E if E happens or fails at S. There are two events, the impossible event ∅ and the sure event Ω, that are determinate in the initial situation Ω and hence never happen or fail.

We call an event not equal to ∅ or Ω a *proper* event. Any proper event E is the disjoint union of the situations at which it happens; its complement is the disjoint union of the situations at which it fails. (In Figure 2, {a,d,e} is equal to {a}∪U and its complement is equal to S∪P.) Thus the situations that resolve E constitute a partition of the entire sample space Ω. Another way to put this is to say that no matter what path we take down the tree—no matter what stop sign we end up in—E is resolved exactly once. We call the set consisting of the situations that resolve E the *resolving partition* for E, and we designate this partition by Ξ_E.

We say that an event E *precedes* an event F if E is determinate in any situation in which F is determinate. If E and F are both proper, then E precedes F if and only if E is always resolved at least as soon as F. In other words, for every path down the tree, the situation that resolves E comes before (or equals) the situation that resolves F. If E is improper (equal to ∅ or Ω), then E precedes any event F.

Variables are defined in the event-tree framework just as they are in the sample-space framework; a variable X is a mapping from the sample space onto another set, say Θ_X. Variables generalize events, for an event can be thought of as a variable with only two values.

A variable X is *determinate* in S if it assigns the same value to all elements of S; otherwise it is *indeterminate* in S. If X is determinate in every situation in which Y is determinate, then we say that X *precedes* Y. We say that S *resolves* X if X is determinate in S but not in S's mother. If S resolves X and X's value for elements of S is x, then we say that S *resolves* X to x.

Any variable X determines two partitions of the sample space Ω, both of which are important: its *value partition* Π_X and its *resolving partition* Ξ_X. The *value partition* consists of events of the form {X=x}. The resolving partition consists of the situations that resolve X. If X is constant, it never resolves, but we call {Ω} its resolving partition.

3 Probability

We assume that a probability has been specified for the outcome of each experiment in our probability tree. This was done implicitly in Figure 1. By specifying probabilities for the outcomes of each experiment, we determine a probability for every event E in every situation S. We write $P_S(E)$ for this probability. The properties of the $P_S(E)$ are summarized by the following axioms.

P1. $0 \le P_S(E) \le 1$ for every situation S and every event E.

P2. If E is impossible in S, $P_S(E) = 0$.

P3. If E is certain in S, $P_S(E) = 1$.

P4. If E and F are incompatible in S, $P_S(E \cup F) = P_S(E) + P_S(F)$.

P5. If S precedes T and T precedes U, $P_S(U) = P_S(T) \, P_T(U)$.

Implications of these axioms have been explored by the author and V.G. Vovk [3, 4, 7].

A variable X has a probability distribution in every situation S; we designate this probability distribution by $P_S{}^X$. Abstractly, $P_S{}^X$ is the function on Θ_X given by $P_S{}^X(x) = P_S(X=x)$.

4 Independence

We say that two events E and F are *independent* if the probability of only one of them can be changed by a given experiment. In other words, for each non-terminal situation S, either $P_T(E) = P_S(E)$ for all daughters T of S, or else $P_T(F) = P_S(F)$ for all daughters T of S. Independence is a causal concept; it says that E and F have no common causes. The experiments that influence one of the events do not influence the other. This is a strong but qualitative condition. It constrains only when probabilities change, not the values they take or the amount by which they change.

Figure 3, in which a fair coin is spun three times, provides the setting for a familiar example of independence. If we let E be event we get heads on the first spin (E={a,b,c,d}) and F the event we get heads on the second spin (F={a,b,e,f}), then E and F are independent. Other examples are provided by the sure event Ω and the impossible event $\varnothing$. Since their probabilities never change, each is independent of any other event.
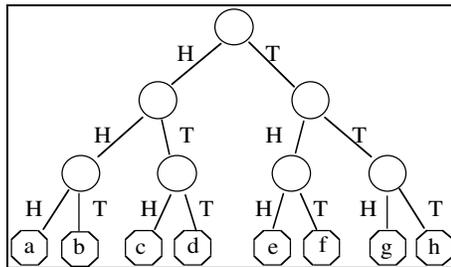
Fig. 3. Three spins of a fair coin.
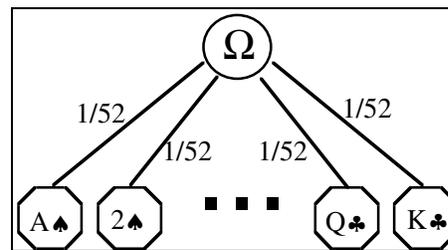
Fig. 4. Choosing one card at random.

We say that E and F are *formally independent* if $P_S(E \cap F) = P_S(E \, P_S(F)$ for every situation S. In other words, E and F are independent in the usual sense with respect to each probability measure P_S. The events E and F in Figure 3 are formally independent as well as independent. Figure 4 gives an example of events E and F that are formally independent but not independent; E is the event we get a spade, and F is the event we get an ace.

The definitions of independence and formal independence extend immediately from events to variables. Two variables X and Y are *independent* if only one of their probability distributions can change in a given experiment. They are *formally independent* if $P_S(X=x \& Y=y) = P_S(X=x)P_S(Y=y)$ for all x and y and every situation S.

The following proposition reveals how independence and formal independence are related: independence is stronger.

> **Proposition 1** If two variables (or events) are independent, then they are formally independent.

In general, we do not stand beside nature and watch how her probabilities change as events unfold in her probability tree, and hence we may be unable to tell whether given variables are independent in her tree. Proposition 1 sometimes gives us grounds, however, for conjecture. If we observe that two variables are independent in the conventional sense in many different situations, then we may conjecture that they are independent in the probability tree sense in nature's probability tree.

In the event-tree framework, as in the sample-space framework, independence can be generalized to conditional independence. But there are several distinct concepts of conditional independence in probability trees. We say that events X and Y are *independent in* S if the probability distribution of only one of them can change in any given experiment at S or below. We say that Y and Z are *independent posterior to* X if they are independent in every situation in which X is determinate—i.e., if the probability distribution of only one of them can change in any experiment after X is resolved. We say that X and Y are *formally independent posterior to* X if $P_S(Y{=}y\&Z{=}z) = P_S(Y{=}y)P_S(Z{=}z)$ holds in every situation S in which X is determinate. Finally, we say that Y and Z are *formally independent given* X if

$$P_S(Y{=}y\&Z{=}z|X{=}x) = P_S(Y{=}y|X{=}x)\,P_S(Z{=}z|X{=}x) \tag{1}$$

holds for every situation S and all values x, y, and z such that $P_S(X{=}x){>}0$. This is stronger than formal independence posterior to X, which only requires that Equation (1) hold in situations in which X is determinate.
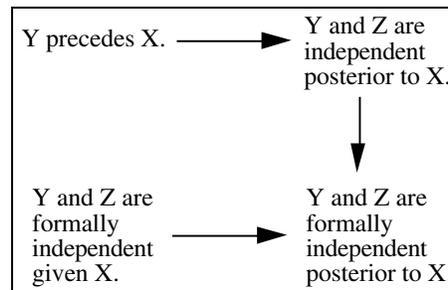
Fig. 5. Logical implications.

Figure 5 summarizes the logical relations among the conditional independence concepts. For the purposes of causal conjecture, Figure 5 is disappointing because there is no arrow from the probability-tree concepts, precedence and independence, to formal independence. Proposition 1 says that when we observe conventional sample-space independence of X and Y (their probabilities multiply), we can explain this observation by the conjecture that X and Y are independent in the probability-tree sense (they have no common causes). Figure 5 does not give us the means to extend this kind of explanation to the conditional case. Since there is no arrow from independence posterior to X to independence given X, independence posterior to X (Y and Z have no common causes after X) will not be an adequate explanation of an observed conditional independence given X.

5 Tracking

Consider two variables X and Z. Suppose that for any possible value x of X, the probability distribution of Z is the same in any two situations that resolve X to x; if S and T are both situations that resolve X to x, then $P_S{}^Z = P_T{}^Z$. Then we say that X *tracks* Z. When X tracks Z, we write $P_{X{=}x}{}^Z$ for the probability distribution of Z in the situations

that resolve X to x. A constant variable tracks every variable and is tracked by every variable.

Coin spinning provides another familiar example of tracking. Suppose we spin a fair coin repeatedly, and let X_i designate the number of heads in the first i spins. Then X_i tracks X_j whenever i<j. Figure 6 illustrates this for the case where i=6 and j=8.

It turns out that when X tracks Z, we can write

$$P_S{}^Z = \sum_{x \in \Theta_X} P_S(X=x)\, P_{X=x}{}^Z \tag{2}$$

in every situation S that resolves X and in every situation S in which X is indeterminate. Since the situation S appears on the right-hand side of Equation (2) only to tell us its probabilities for X, we may say that the probabilities for Z are initially governed, in a certain sense, by the probabilities of X. Of course, after X is resolved, the equation no longer applies, and the probabilities of Z are left to evolve on their own. In Figure 6, for example, the probabilities for X_8 are tied to those for X_6 only through the first six spins; after that X_6 is settled and X_8 is on its own.

The message of Equation (2) can also be expressed using the language of causality. It says that until is X settled, Z has no causes distinct, in identity or magnitude, from X's causes.
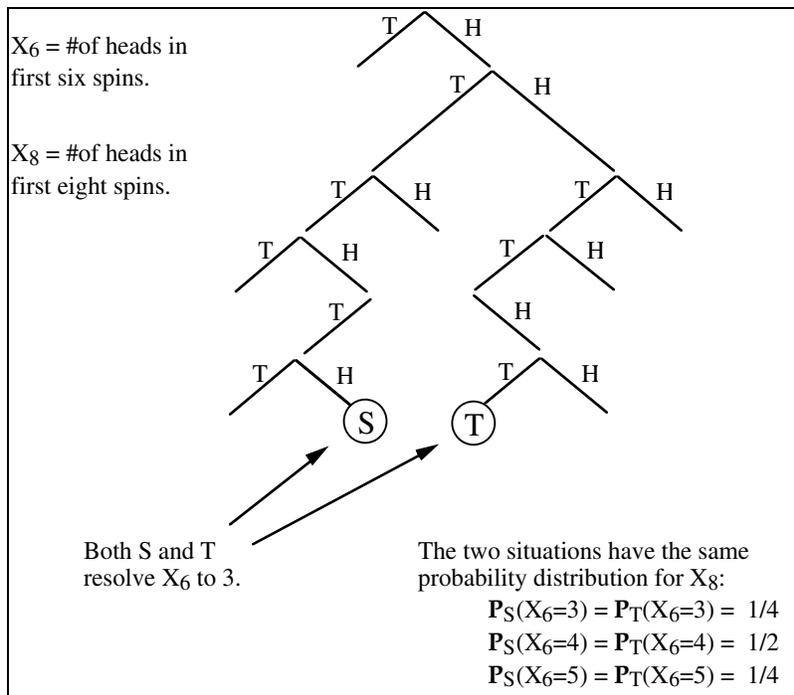
Fig. 6. Fragment of a probability tree for coin spinning.

Although Figure 6 is a natural example of tracking, the definition of "X tracks Z" does not require that X precede Z. It is possible that Z might instead actually precede X. An example of this is provided by Figure 7. Here we have three piles of cards; the first contains the deuce and three of spaces, the second contains the four and five of hearts, and the third contains the six and seven of hearts. We choose a pile at random and then choose a card at random from the pile. The variable X is the number on the card chosen,

while Z is the suit. Since every situation that resolves X resolves it to a different value, X strongly tracks every other variable, including Z. But Z obviously precedes X.
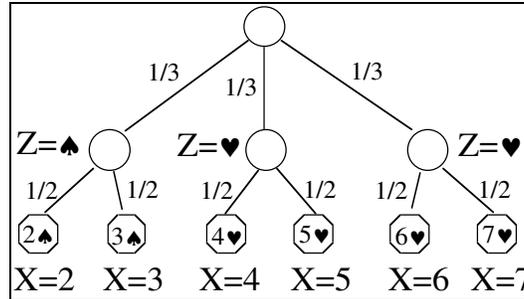

Fig. 7. An example where X tracks Z but Z precedes X.

The example in Figure 7 has another interesting feature: Z is a function of X (if X is 2 or 3, then Z is spades; if X is 4, 5, 6, or 7, then Z is hearts). This feature is not accidental, as the following proposition makes clear.

> **Proposition 2.** The variable Z is a function of the variable X if and only if Z precedes X and X tracks Z.

A certain clash of intuitions is evoked by the statement that a function of X must precede X. When Z is a function of X, we often think of using X to find Z, and in this picture, X comes first. In order to avoid confusion, we must bear in mind the distinction between the observer in the probability tree, who goes down the tree finding out events as they happen, and a person outside the tree, who uses values of certain variables to compute values of others. The second person may also be an observer in a probability tree, but it will be a different probability tree.

In general, we cannot tell for certain whether a given variable X tracks another variable Z in nature's probability tree. But the following theorem allows us to sometimes conjecture that it does.

> **Theorem 1.** If X tracks Z and S is a situation in or before X's resolving cut such that $P_S(X=x) > 0$, then $P_S(Z=z|X=x) = P_{X=x}^Z(z)$.

This theorem tells us that X tracking Z can serve as a causal explanation of the observation that the probability distributions of Z given X are the same in many different situations.
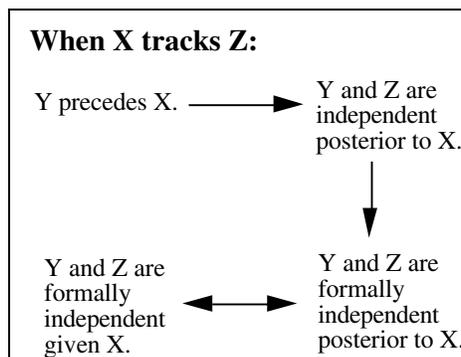

Fig. 8. Tracking and conditional independence.

The following theorem indicates another way in which tracking can serve as part of a causal explanation.

> **Theorem 2.** If X tracks Z and Y and Z are formally independent posterior to X, then Y and Z are formally independent given X.

Figure 8 adds Theorem 2 to Figure 5. As this figure indicates, X's tracking Z together with the independence of Y and Z posterior to X imply the formal independence of Y and Z given X. In other words, tracking together with posterior independence can provide a causal explanation of observed sample-space conditional independence.

6   A Causal Interpretation of Bayes Nets

Recall that a Bayes net is a directed acyclic graph with distinct variables from a probability space (P,Ω) as nodes, with the property that each variable is independent, given its parents, of its nondescendants.

Theorem 2 provides one way to give causal explanations for the conditional independencies in a Bayes net. In order to express this formally, let us say that a directed acyclic graph with variables as nodes is a *causal net in the tracking sense* if in nature's probability tree,

- each variable is tracked by its parents, and
- each variable is independent, posterior to its parents, of its nondescendants.

Theorem 2 tells us that a causal net in the tracking sense is a Bayes net in the conventional sense with respect to the probability distribution for every situation.

This result provides a framework within which users of Bayes nets can express, elaborate, examine, and defend their causal assumptions. Three points should be kept in mind in this process.

- Saying that a Bayes net is causal net in the tracking sense does not necessarily mean that X is settled before Y when there is an arrow or a directed path from X to Y. If this is intended, it should be made explicit.
- In many cases, the claim that a Bayes net is causal net in the tracking sense is not plausible.
- Causality in the tracking sense is not the only possible causal interpretation for Bayes nets. But whatever interpretation is used, it should be explicitly stated and examined.
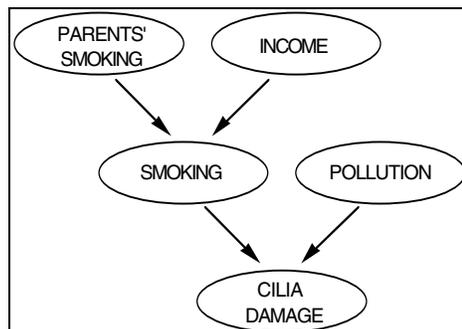

Fig. 9. A hypothetical Bayes net.   Fig. 10. An elaboration.
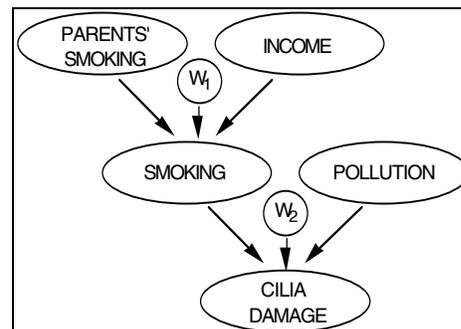
Figure 9, adapted from Spirtes [5], can illustrate the clarification that is possible when we undertake to be explicit about the causal interpretation of a Bayes net. In this purely hypothetical example, parents' smoking and the person's income are said to be "causes" of her smoking status, and this status together with the level of air pollution where she lives are said to be "causes" of the extent of damage to the cilia in her lungs. A little thought shows, however, that these assertions cannot be plausibly interpreted in terms of tracking and posterior independence. At the point in time where a person's income and

her parent's smoking habits have been determined, there are surely other factors that have also already come into play in influencing her smoking status.

One way to get a causal interpretation of Figure 9 might be to enlarge it, as in Figure 10. Here W_1 represents the other factors influencing smoking status by the time its parents in Figure 9 are resolved, and W_2 similarly represents the other factors influencing cilia damage. If we are willing to believe the additional independencies indicated by Figure 10, then it will be a causal net in the tracking sense, and since Figure 10 being a Bayes net implies Figure 9 is also a Bayes net, the causal explanation of Figure 10 can be taken as a causal explanation of Figure 9.

This example also illustrates the point that tracking provides only one causal explanation of Bayes's nets, and that others are often needed. Since the damage that smoking does to the lungs is progressive, it may be more useful to think of a causal interpretation in which each additional act of smoking merely proportionally raises the expected level of lung damage. This leads to the concept of linear sign, which is weaker than tracking. There are results for linear sign analogous to Theorems 1 and 2:

- if X is a linear sign of Z, then the linear regression coefficient of Z on X is the same in every situation, and
- if X is a linear sign of Z, and Y and Z do not change in mean when X does not change in mean, then Y and Z have zero partial correlation linearly accounting for X in every situation.

7   Conclusion

This paper has demonstrated one way in which a Bayes net can arise from qualitative and overtly causal conditions in a probability tree. The demonstration should be taken only as illustrative, for many Bayes nets are not causal nets in the tracking sense, and there are other conditions on a probability tree that can lead to the conditional independence relations that define Bayes nets.

The existence of precise causal interpretations for Bayes nets raises a challenge to anyone who wishes to attribute causality to any Bayes net—whether the person be a social scientist working with a Bayes net inferred from data or an AI engineer constructing a Bayes net for use in decision making. It is no longer enough to say vaguely that the Bayes net is causal; instead, one must provide a precise conjecture about the causal conditions that might give rise to the Bayes net.

The ideas developed here go beyond networks encoding conditional independence relations to similar directed acyclic graphs that encode partial uncorrelatedness or unpredictability relations. This includes the structural equation models used in the social sciences. In these cases, too, we can give causal conditions in terms of an underlying probability tree that will explain the observed or conjectured conditional independence, conditional unpredictability, or partial uncorrelatedness relations, and before we accept a causal claim, we will want to ask that it be made precise in probability-tree terms.

Acknowledgment

References

1. Kenneth A. Bollen: Structural Equations with Latent Variables. New York: Wiley. 1988.
2. Pearl, J. Probabilistic Reasoning in Intelligent Systems. San Mateo, California: Morgan Kaufmann 1988.

3. Shafer, Glenn: Can the various meanings of probability be reconciled? In: G. Keren and C. Lewis (eds.): A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues. Hillsdale, New Jersey: Lawrence Erlbaum 1993, pp. 165-196.

4. Shafer, Glenn: The Art of Causal Conjecture. Cambridge: MIT Press 1995.

5. Spirtes, Peter, Clark Glymour, and Richard Scheines. Causation, Prediction, and Search. Lecture Notes in Statistics 81. New York: Springer 1993.

6. D.J. Spiegelhalter, A.P. Dawid, S. L. Laurtizen, and R. G. Cowell: Bayesian analysis in expert systems (with discussion). Statistical Science, 8 219-283 (1993).

7. Vovk, V.G. : The logic of probability. Journal of the Royal Statistical Society, Series B, 55 317-351 (1993).