The Situation of Causality

Glenn Shafer

gshafer@andromeda.rutgers.edu

5 May 1996

Causality in the abstract is a grand theme. We take it up when we want to penetrate to the bottom of things—to understand general laws that govern the working at the world of the deepest and most detailed level.

In this essay, I argue for a more situated understanding of causality. To counter our desire for ever greater generality, I suggest that causal relations, even those that hold only on average, require context. To counter our desire for ever greater detail, I suggest that causal relations may exist only at a certain level of granularity.

The case for the situatedness of causality is based on the epistemic character of causality. A causal structure is a structure for predictions that might be made by an ideally knowledgeable and observant scientist. It tells us about the unfolding of that scientist's knowledge. It has objective aspects, because knowledge must have objects. But it is also subjective aspects, because knowledge must also have a subject, and possibilities and probabilities that are meaningless without reference to that subject.

The epistemic character and situatedness of causality tell us something about what kinds of objects should be called causes. The objects available to statisticians—variables and Moivrean events—are not causes because they are not situated. But they can be used to detect and aggregate causes.

The arguments advanced here build on my forthcoming book, *The Art of Causal Conjecture* (Shafer 1996), which develops a detailed account of causal relations in probability trees.

1 Using Probability Trees to Talk About Causality

The probability trees in Figures 1 and 2 can be understood as situated causal structures. They will serve to anchor our discussion of the situatedness of causality.

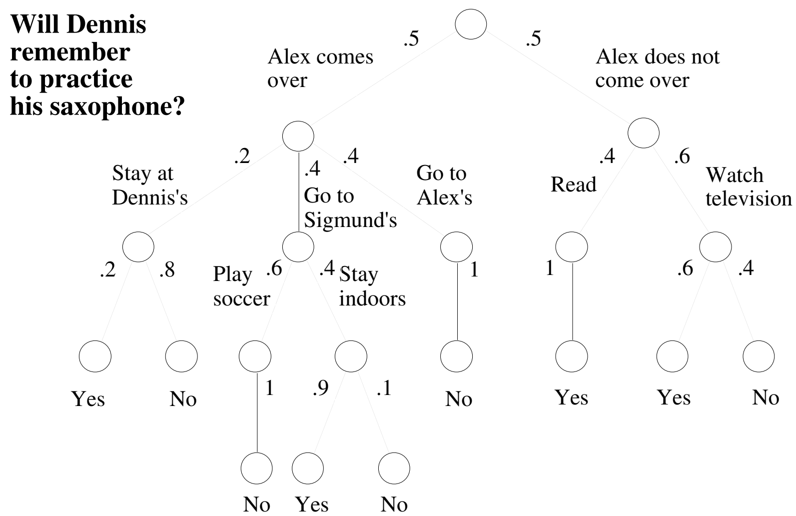


Figure 1 The habits of Dennis and his friends on summer afternoons. The numbers are probabilities for what will happen in each situation (circle). Whether Dennis remembers to practice is influenced by his and his friends' other activities. For example, he is certain to practice if he reads.

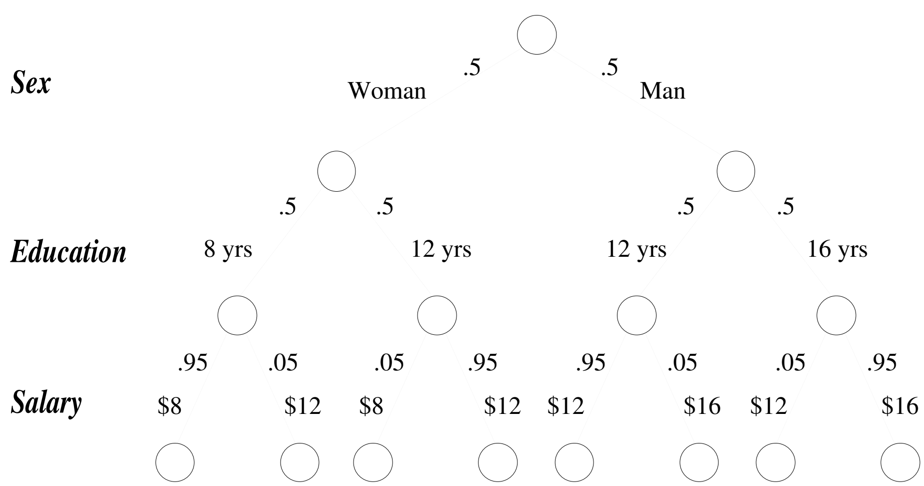


Figure 2 The norms of an imaginary discriminatory society. This society provides more education to its men than to its women, although there is

> some overlap. Both men and women are generally paid in proportion to their education. Employers may deviate from this proportionality for exceptionally capable or hapless employees, provided they stay within the range of pay customary for the employee's sex.
>
> Most of the people who have 12 years of education are paid $12. But a few women are paid less and a few men more. Hence women with 12 years of education are paid less on average than men with the same education. Similarly, most of the people who are paid $12 have 12 years of education, but this group includes a few men with 16 years of education and a few women with only 8. Hence women who are paid $12 have less education on average than men paid the same.[1]

Probability trees provide a framework for a variety of causal talk. We can talk about *singular causation*—why Dennis forgot to practice his saxophone on a particular occasion or why a particular man is well paid. And we can talk about *average causation*—what makes Dennis likely to practice or why women are poorly paid on average.

We can also distinguish between contingent and structural causes. A step or sequence of steps in a tree can be called a *contingent cause*; Dennis's going to Alex's house is a contingent cause of his forgetting to practice. The habits, preferences, norms, or laws of physics that determine the possibilities and their probabilities can be called *structural causes*; Dennis fails to practice at Alex's house because he prefers to do other things there.

---

[1] Situations in the United States where discrimination against women is alleged often have similar statistical features. Women are paid less on average than men with the same education but have less education on average than men with the same pay. See Finkelstein and Levin (1990) and Dempster (1988).

It is appropriate to call steps in a probability tree contingent because alternatives are spelled out. Instead of going to Alex's, the boys might go to Sigmund's. Structural causes, in contrast, are constant within the tree. We can imagine Dennis having different preferences at Alex's house, but this lies outside the story, and alternative possible preferences are not specified.

When young people are tutored in how their society works, contingent causes—causes with alternatives actualized in the society—are often emphasized. In the society of Figure 2, one might say that a particular person has 12 years of education because she is a woman and was among those girls inclined to stay in school. She is paid $12 because she is a woman with 12 years of education and performs adequately as an employee. These kinds of statements are easily understood by anyone familiar with the norms of the society.

From an external perspective, structural causes are more salient. From outside the society of Figure 2, we realize that the causes of a person's salary include not only the contingencies permitted by the norms of the society but also those norms themselves. The woman is paid $12 because the society deems this appropriate for a woman with 12 years of education.

While acknowledging the validity of causal talk associated with pictures like Figures 1 and 2, we are inclined to feel that this talk does not go very deep. These pictures are coarse and apply only in very special circumstances. The are very situated. Surely a genuine understanding of causality requires a more detailed look at behavior and a broader scope. Surely it requires us to speak of fundamental rules or mechanisms that apply more widely than to a single child or a single society.

But once the issue is faced squarely, it also appears reasonable to ask what grounds there are for thinking that causal explanations can always be made more detailed or given broader scope. Is there is an ultimate level of causal explanation that is not both granular and situated?

2 The Epistemic Character of Possibility and Probability

In order to discuss the potential for refining and broadening causal structures represented by probability trees, we need to understand the meaning of the possibilities and probabilities in these trees. In my view, their meaning is more epistemic than ontological, and hence, in a certain sense, more subjective than objective.

According to Figure 2, a woman with eight years of education has the possibility, at the point where it is settled that she will have only that much schooling, of earning either $8 or $12. (If this were not always true—if, for example, the possibility of earning $12 were already closed for some women at this point—then we would not agree to call the tree a causal structure.) What does this mean? Without entering into a general discussion of the meaning of possibility, let me suggest that in this context possibility comes down to the impossibility of prediction. When we say that either salary is possible for the woman, we mean that no one, not even an ideal observer—a demigod or a highly astute and well informed sociologist—can say which she will receive.

This concept of possibility is epistemic. The observer is ideal, to be sure, but our reference to the observer means that we concerned with knowledge and observation, not with things in themselves. When we say that it is equally possible that a woman will earn $8 or $12, we are not denying that God has already foreseen and even ordained what she will earn. But if God has already ordained the amount of her salary, then he has also arranged that no one can guess the amount.

A similar epistemic interpretation can be given to the probabilities in a causal structure. Figure 1 gives odds 6 to 4 in favor to Dennis watching television rather than reading on a given afternoon without Alex, and these odds, like the possibilities themselves, gain their meaning from the impossibility of prediction. Without God's foreknowledge, neither we nor an ideal observer has any way of selecting bets at these odds that will do better than break even in the long run. All we know is that Dennis will read 40% of the time and watch television 60% of the time.

In many discussions of causality, especially in statistics and artificial intelligence, we find a more ontological picture of randomness. God is left out of the picture, and Nature is portrayed as a mover of events. Nature determines what will happen next by rolling dice. We can appropriate this imagery for our more epistemic account, but with a clarification. Nature does not determine what will happen next. Nature merely observes how the dice fall. Nature is our ideal observer—an observer who sees and knows more than any real scientist but still can offer only odds rather than categorical predictions. If we adopt this way of talking about Nature, then we can say that a probability tree is causal precisely when it is Nature's probability tree—when its possibilities and probabilities are Nature's.

The epistemic interpretation of possibility and probability in Nature's probability tree has both subjective and objective aspects. Possibility and probability are always subjective, because they depend on the position of an observer.[2] When the observer is an idealization that represents the limit of knowledge and experience attainable by actual observers, these possibilities and probabilities are also objective, because they are borne out by experience. Actual observers approximately break even in the long run when betting at Nature's odds, no matter what their strategy for placing bets.

The objective aspect of probability in Nature' s tree is most salient and familiar when the same causal structure is repeated exactly many times, as in Figure 3, for then the most obvious strategy for placing bets is to make the same bet over and over, and the fact that this strategy breaks even means that each probability in the repeated causal structures is a frequency. But exact repetition is not required. Provided only that there are many opportunities to bet on every path through the tree, we can formulate the assumption that

---

[2] This modest conception of subjectivity does not have the psychological element that has dominated discussions of subjective probability since the work of David Hume in the eighteenth century. But it is close to the seventeenth-century thinking of Jacob Bernoulli (Shafer 1996: Chapter 1).

any strategy Nature can think of will approximately break even, and this assumption will have an implication for frequencies: the average frequency with which we win our bet about what step is taken next will approximately equal the average probability of the steps bet on (Shafer 1996: Chapter 4).

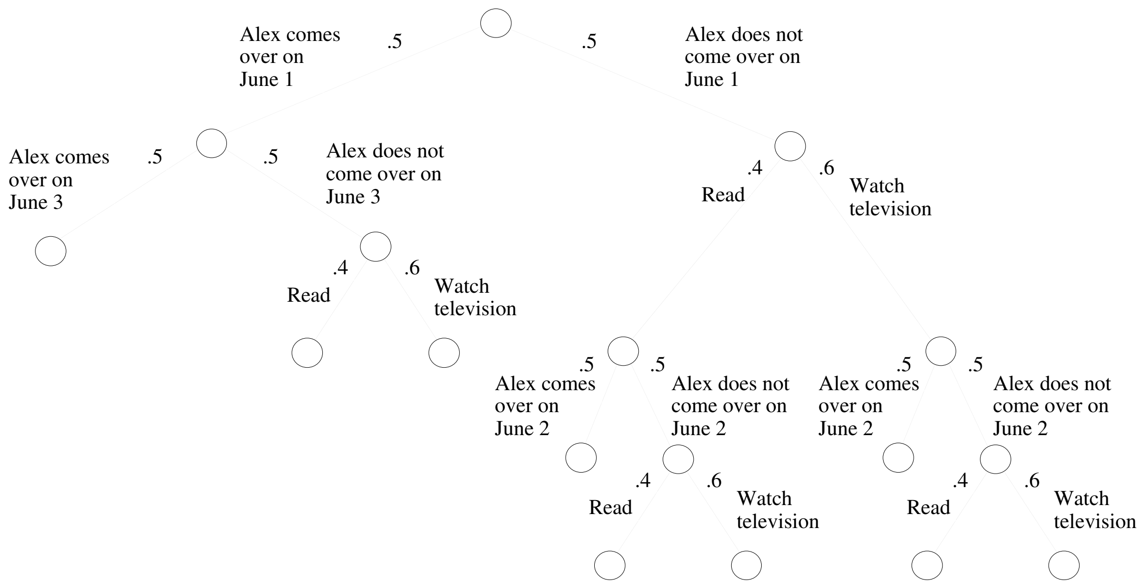


Figure 3 Dennis and Alex on successive summer afternoons. Here we see a fragment of Figure 1 happening twice, first on June 1 and then either on June 2 or June 3. The reader may imagine a much larger tree in which this fragment (and even the whole tree of Figure 1) is repeated many times along every path.

3 Refinement and its Limits

It is always mathematically possible to refine a probability tree. Figure 4 illustrates how we might do this for the branch of Figure 2 that applies to women. It can never be taken for granted, however, that a causal probability tree has a causal refinement. A causal probability tree asserts a special form of knowledge for Nature, and the fact that Nature's knowledge has this form at one level of granularity does not imply the same at another level of granularity.

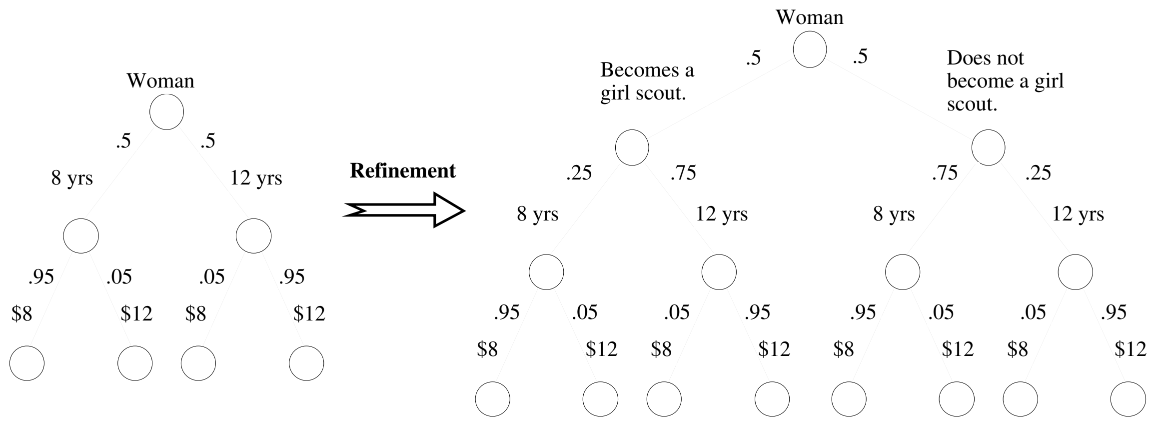


Figure 4 Some detail about how the educational level of a woman is determined. The more refined tree does not contradict any causal assertions in the original tree. It continues to assert, for example, that a woman with eight years of education has only a 5% chance of earning $12. How the woman came to have only eight years of education does not matter.

If the probability tree with which we begin is thought of as a structure repeated many times as Nature moves through her tree, then refinement may fail simply because different instances refine in different ways. It may be that each woman in the society of Figure 2 is influenced by different events in her childhood. For one woman, becoming a girl scout be influential, as in Figure 4. For another, whether an older brother survives to adulthood may be crucial, etc. Thus there might be a refinement for each woman, but no refinement applicable to everyone.

Refinement may also fail in a more fundamental way. It may be that there are no hints in the childhood or background of any of the women that enables Nature to predict in advance whether they will continue in school after the eighth grade. Or it may be that while there are some hints, Nature is not able to make out a stable pattern that can provide a basis for improving her probabilistic prediction.

4 Broadening and its Limits

A probability structure can be broadened as well as refined; Figure 5 gives an example. Like refinement, broadening cannot be taken for granted. From the assumption that Nature can make probabilistic predictions under one set of conditions, it does not follow that she can do so under another set of conditions.

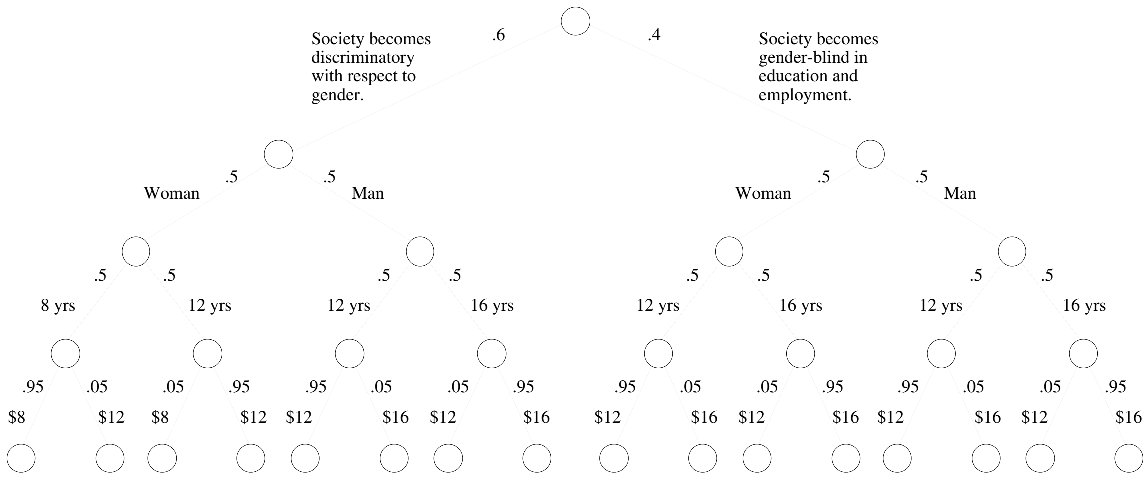


Figure 5 The probability tree of Figure 2 is the left branch of this broader tree. The right branch describes another possibility for how the society might evolve.

Much of the work of scientists and engineers is devoted to the design of conditions—in laboratories and factories—under which actions have predictable consequences. As Cartwright (1989) has observed, this effort itself is evidence that the world outside the laboratory and factory is more chaotic. Sometimes, of course, we can construct different laboratories that work in different ways. This will certainly give us a broader view of the possibilities in Nature, and it may suggest conjectures about causal relations in Nature outside the laboratory. But sometimes these conjectures are very limited in their breadth.

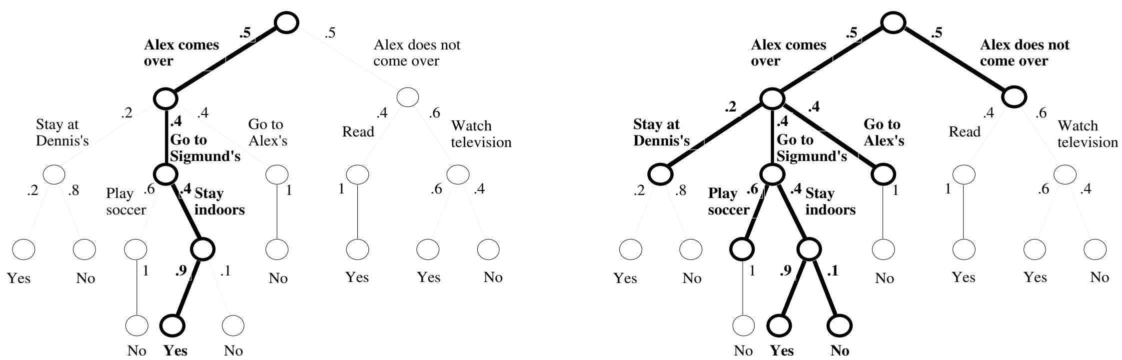
5 Causal Law as Constitutive Rule

A probability tree is a thoroughly compiled representation of a causal structure. It spells out completely all possible courses of events and their probabilities. We might imagine that Nature has also a more compact representation, consisting of causal laws—rules that lay out, for situations of various types, the possibilities for the next step and their probabilities. These laws can make the probabilities and possibilities depend on the properties of various objects in the situations. In Figure 1, for example, Nature might have a law that expresses the probabilities for a child's reading or watching television as a function of properties of that child.

A causal law can be valid, and can be borne out by experience, even though it produces different probabilities and possibilities every time Nature uses it. It is borne out by experience if Nature can find no way, as she applies the law over and over, of ruling out any of the possibilities the law allows, and she can find no strategy for placing successive small bets at the odds given by the law that does not approximately break even. The fact that Nature can use a flexible law, applicable to a variety of situations, reinforces the point, made at the end of Section 2, that the objective aspect of possibility and probability does not depend on exact repetition of given possibilities and probabilities.

If Nature has laws that apply broadly across her tree, and hence provide a compact representation of that tree, we may wish to regard these laws as more fundamental than the tree itself and hence as the proper object of causal investigation. There is scant reason to assume, however, that laws for the branching possibilities and probabilities in Nature's tree are unique. In the society of Figure 2, for example, we might choose different aspects of the difference between men and women as the basis a law predicting education and salary. If these different laws produce equally good predictions, and there is no broader causal tree that picks out one of them (i.e., no broader set of societies where the laws diverge and only one agrees with Nature), then we may have no way of saying which of the laws is fundamental, and Nature's tree itself will seem more fundamental than any of them.

We should also acknowledge that part of Nature's tree has a greater claim to reality than the rest: the path actually taken as events unfold. The regularities that Nature learns lend immediate meaning to a slightly wider swath as well: the immediate predictions Nature makes as she moves along the path (see Figure 6). It is here, perhaps, that we should stage the debate sketched in the preceding paragraph. Some will argue that the predictions in this swath along Nature's path are less fundamental than causal laws that produce these predictions, while I am suggesting that more than one set of laws might produce them.[3]



The actual course of events

Predictions validated by a longer course of events

Figure 6 The fundamental aspects of Nature's tree. The path actually taken is most fundamental. The immediate possibilities and probabilities recognized by Nature as this path is taken come next, because they represent the longer run of experience that validates Nature's rules, and perhaps also the experience that Nature uses to formulate these rules. The

---

[3] These issues are discussed by Dawid (1984) in a slightly different framework, in which experts engage in sequential prediction. Our framework does involve at least one complication not found in sequential prediction: because we allow refinement, the swath along Nature's path may fail to be well defined.

> tree as a whole has a secondary reality, for it represents an extrapolation of Nature's rules.

Figure 6 raises another interesting point. Perhaps there are different ways of formulating rules—causal laws—that produce the same predictions down the swath actually taken but different predictions in the tree as a whole. In other words, they produce different trees. If these different trees really do lie outside Nature's total experience, then there seems no grounds to choose between the trees or the laws that produce them. In saying this, however, we must remember that Nature's experience is very broad. Many of the causal debates among us mere mortals reflect real differences of opinion—differences about how our limited experience extrapolates to Nature's broader experience.

Another familiar thread of thought enters here: perhaps we can choose among different causal laws leading to the same predictions on the basis of their complexity. The principle that simpler laws should be preferred seems reasonable, we should take it with a dose of salt. There is no reason to suppose that complexity can always be assessed precisely enough to force a choice among causal laws. If we are content with the probability tree itself (or with the swath along the true path) as the fundamental representation of causality, then this gives no occasion for unhappiness.

A similar attitude can be taken towards the idea of using the success of causal laws to identify natural properties. A property of objects (or even a way of individuating objects) is meaningful and fundamental precisely to the extent that it enters into simple causal laws. This is a useful principle, but there is no reason to assume that it will not leave wide latitude for judgment and taste.

7 Causal Law as Rule of Construction

When we think of a causal law as a rule for specifying possibilities and probabilities in situations in Nature's tree, we are giving the law take a double role. It tells what "action" or "experiment" will be performed in each situation, and it tells the possible outcomes and their probabilities. It is sometimes useful to separate these roles. This is particularly appropriate when we want to claim the right to choose the action, as in decision theory (in statistics) and planning (in artificial intelligence). In this case, the causal law tells not what action will be taken but what actions can be taken. It is a constraint on the construction of a probability tree rather than a rule constitutive of a probability tree.

Our emphasis in this paper is on the working of a given machine. We have suggested that causality is a matter of understanding the possibilities within a psychological machine like Figure 1 or a socio-economic machine like Figure 2. When we turn to rules for constructing such machines, we seem to have climbed to a higher vantage point.

On the other hand, rules of construction for one machine can be thought of as constitutive rules for a larger machine. And this returns us to the message of Section 4: the larger machine, since it describes a broader part of the world, may be more interesting, but this does not mean that it has the regularity or stability that will merit its being treated as a causal structure. Nature may not be able to formulate rules for predicting possibilities and probabilities in this broader part of the world.

One compromise is of particular interest. If we are indeed choosing actions, then it is natural to list the alternatives without giving probabilities for the choice, and this produces what has traditionally been called a decision tree (Raiffa 1968). This is illustrated by Figure 7. The mathematical theory of probability trees generalizes readily to decision trees (Shafer 1996: Chapter 12).

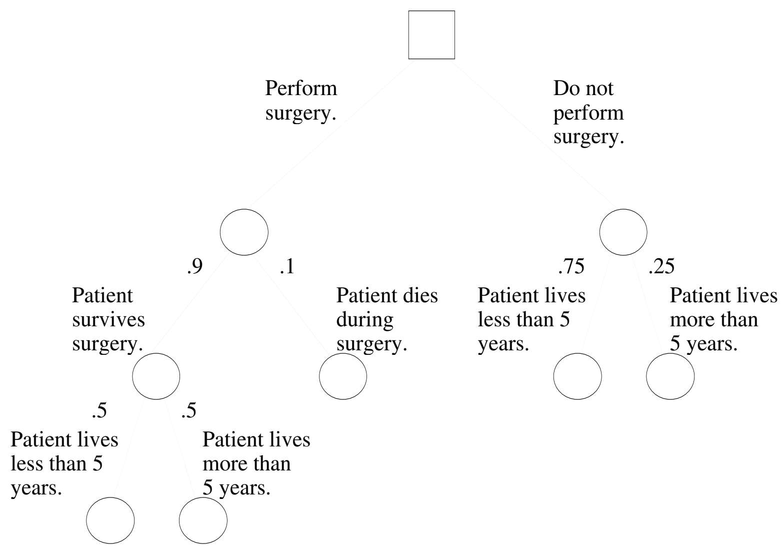


Figure 7 The square is a "decision node"—a situation in which a decision is taken. The choices in the decision are not given probabilities.

While the decision-tree generalization of probability trees is adequate for an abstract understanding of causality when actions are allowed, the practical implementation of the decision or planning process might benefit from these trees being described in terms of rules of construction.

8 Are Variables Causes?

It is common, in applied statistics, to speak of variables as causes. The idea that smoking causes lung cancer, for example, is often expressed by saying that the total number of cigarettes smoked (a variable) causes lung cancer. Let me explain how this practice arises and why it can be misleading.

What is a variable? Statisticians think of a variable as a measurement made on an individual—something that varies from individual to individual. For probabilists, on the other hand, a variable is a function on the sample space. The sample space for a probability tree is simply the set of paths through the tree, and hence a variable is anything determined by the path. Figure 8 shows an example.

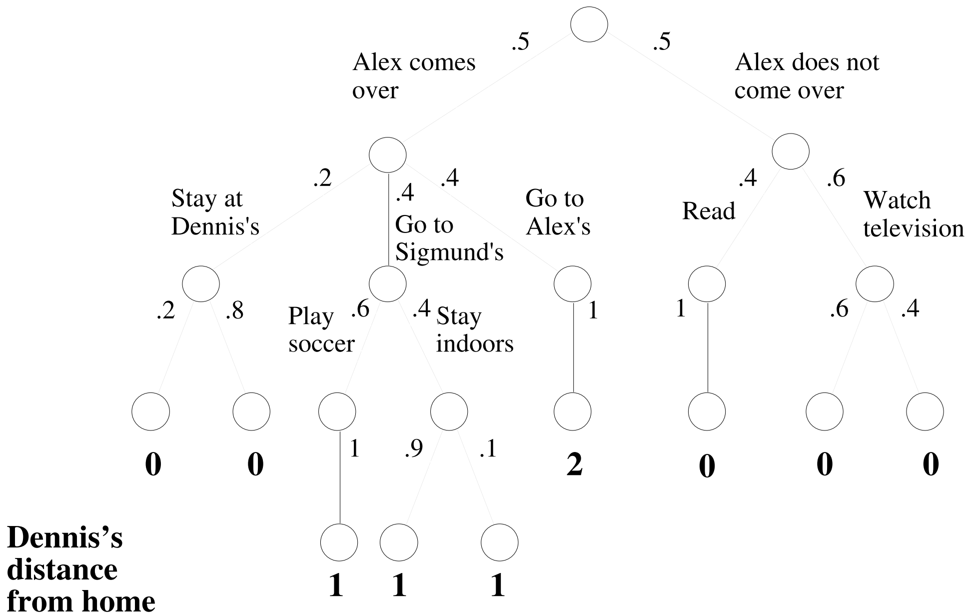
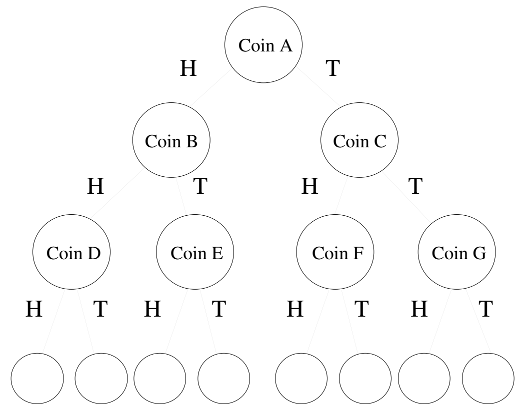


Figure 8 At the end of the afternoon, Dennis will be at Alex's house (2 blocks away), at Sigmund's house (1 block away), or at home (0 blocks away). His distance from home is a variable.

To get from the statistician's conception of a variable to the probabilist's conception, we must identify individuals in our probability tree. Individuals appear in situations, and the measurements we take on these individuals are sometimes represented by steps in the tree. In Figure 9, for example, we can identify six balanced coins, each of which is spun in the situation where it appears. The result, heads or tail, is a variable, since it varies from coin to coin. But it is not probabilist's variable, since it is not something that is determined on every path through the tree. To get variables in the probabilist's sense, we conflate coins in situations across the tree, obtaining a coin that appears no matter what path is taken. We conflate coins B and C, and we conflate coins D, E, F, and G, thus reducing the six coins to three: the "first coin," the "second coin," and the "third coin." We then have three variables in the probabilist's sense, X_1, X_2, and X_3. These three variables correspond to the statistician's single variable—how a coin falls.



X_1	1	1	1	1	0	0	0	0
X_2	1	1	0	0	1	1	0	0
X_3	1	0	1	0	1	0	1	0
Y	3	2	2	1	2	1	1	0

X_1 = first coin falls heads
X_2 = second coin falls heads
X_3 = third coin falls heads
Y = Total number of heads

Figure 9 Nature encounters three fair coins along every path. Each coin is spun, and the number of heads is measured by X_1, X_2, X_3, and Y.

Suppose I bet on heads on each spin in Figure 9, and I use my winnings to invest in real estate. Is it then colloquial or reasonable to call the variables X_1, X_2, and X_3 causes of my eventual wealth? Is it reasonable to call their sum, Y, a cause of my eventual wealth? These locutions do seem reasonable, but perhaps only as abbreviations for more careful statements. The steps in the tree, where the coin falls heads or tails, are indeed causes of my eventual wealth. The variables are summaries of these causes.

We can think about smoking and lung cancer in a similar way. The steps in Nature's tree where I smoke a cigarette—or where things happen that increase my likelihood of smoking—are genuine causes of lung cancer. The total number of cigarettes I smoke measures these causes—it measures, as it were, how many of them occur.

Of course, steps in Nature's tree can be summarized in many different ways. It is misleading, therefore, to treat a variable as anything other than a convenient summary. This is especially true when more than one variable is considered. Different variables can sometimes summarize the same steps in Nature's tree, and if the variables, instead of the steps, are given prominence, we can fall into misguided debates about which variables are real causes.

If we believe that Nature's tree is generated by uniquely defined causal laws, and we believe that these causal laws can be stated in terms of properties of individuals in the situations where they apply, then we might well call the variables specifying these properties causes. The arguments made in Section 5 give us good reason not to take the first step on this road; the tree may be more fundamental than any rules generating it. What we have just learned about variables in probability trees gives us good reason not to take the second step. To obtain a variable in a probability tree, we must, in general, conflate individuals across the tree, and this means conflating properties in ways that may, in fact, be quite arbitrary. This arbitrary choices that must be made contribute to the meaning of any variable used to describe these properties and thus makes it inappropriate to regard these variables as fundamental.

The arbitrariness involved in conflating individuals is not evident in Figure 9, for there we begin with the idea that "heads" is well-defined for all six coins. But many nations have coins without a person's portrait on one side, and if such coins are among the six shown in the figure, we must decide what side to call heads. If we are concerned about measuring the causes of my wealth, we might choose the side I bet on. But this will hardly produce a variable that is meaningful as a basic ingredient of Nature's causal laws.

Similar problems arise when we compare individuals across Nature's tree who differ because of steps taken earlier in the tree. An earlier step may give an individual properties that do not even apply to individuals on other branches. For example, an individual who has had surgery to remove an organ will have properties that relate to that surgery (e.g., how much blood was lost) and will no longer have properties related to that organ (e.g., its temperature). If we regard one of these properties as causally important for the individuals that have it, and we want to express that property as a variable, then we must decide how it is defined for individuals across the tree. We must decide, for example, on an organ temperature for a person who does not have the organ.

9 Humean and Moivrean Events

While variables are the objects most commonly called causes in applied statistics, it is at lest as common in the philosophical literature to call events causes. (See, for example, Mackie 1974.) Is this appropriate?

In this section, I distinguish between two types of event. On the one hand, a step in the tree—or a set of steps—is a *Humean event*. On the other hand, a set of paths through the tree is a *Moivrean event*. Humean events correspond to the intuitions of philosophers and can correctly be called causes. Moivrean events, although they are what probabilists call events, should not be called causes.

I have already argued at length that a single step in Nature's tree can correctly be called a cause. But as Figure 10 illustrates, a single step at one level of description becomes something more complicated at a more refined level. I call any set of steps in Nature's tree that appears as a single step at some level of description a *Humean event.*

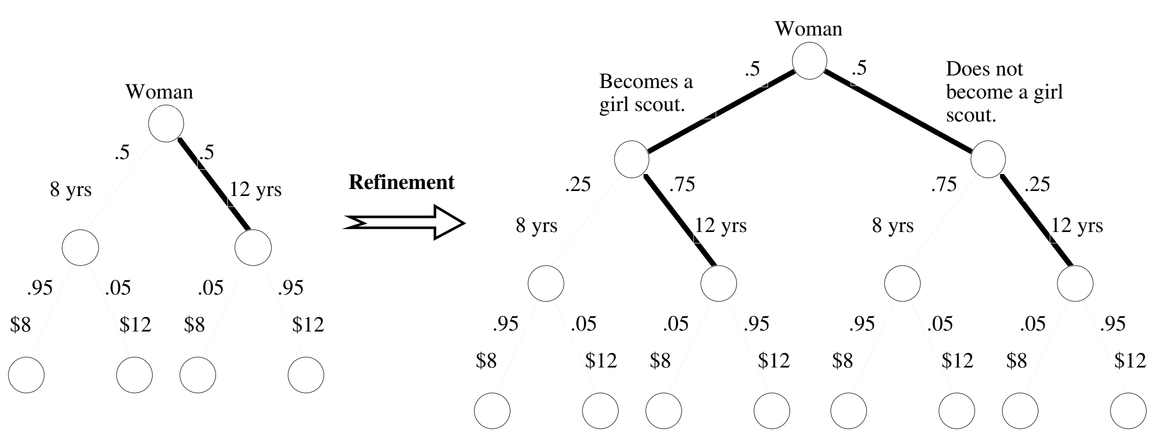
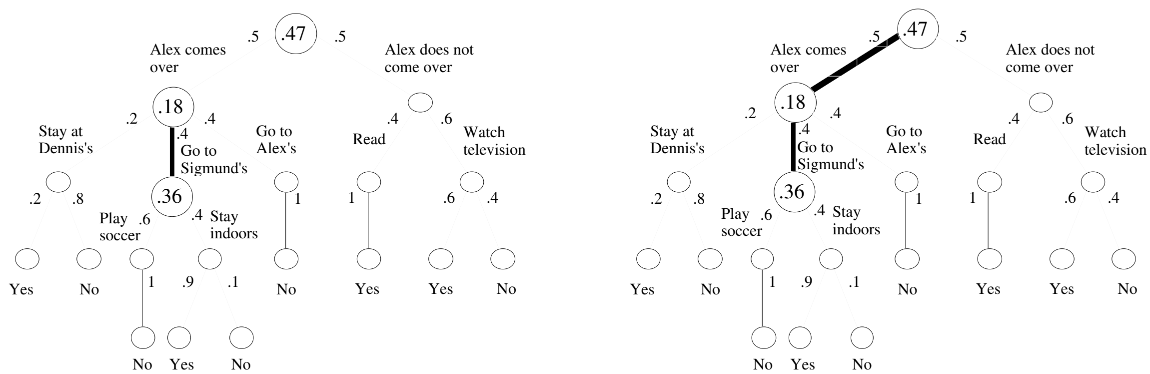


Figure 10 This second look at the refinement of Figure 4 shows how a single step in the original tree becomes four steps in the refinement. The four steps form two chains, both of which begin with the person's gender

being determined as "woman" and end with her education being determined as "12 years."

A Humean event can properly be called a cause precisely because it is situated. It has a starting point (a context) as well as an ending point. When we call an event a cause, we always have such a starting point in mind, even if we do not specify it explicitly. If, for example, we speak of Dennis's going to Sigmund's house as a cause of his remembering to practice, then we have in mind one of the two starting points shown in Figure 11. The difference is important, because the event on the left raises the probability of Dennis's remembering to practice, while the event on the right lowers it.



Going to Sigmund's house, after Alex has come over.

Going to Sigmund's house, before Alex has come over.

Figure 11 Two Humean events that end by Dennis's going to Sigmund's house. The Humean event that begins after Alex has come over and ends at Sigmund's house raises the probability of Dennis's remembering to practice from 18% to 36%. The Humean event that begins before Alex comes over and ends at Sigmund's house lowers the probability that Dennis remembers to practice from 47% to 36%.

Since the publication of Abraham De Moivre's *Doctrine of Chances* in 1718, the word "event" has been used within probability theory in a way that is not situated. A Moivrean event is a subset of the sample space, just as a variable is a function on the sample space. (Thus it is equivalent to a variable that takes only two values—say the value one on the elements of the sample space in the subset and the value zero on elements outside the subset.) In a probability tree, the sample space is the set of paths through the tree, and hence a Moivrean event is a subset of these paths.

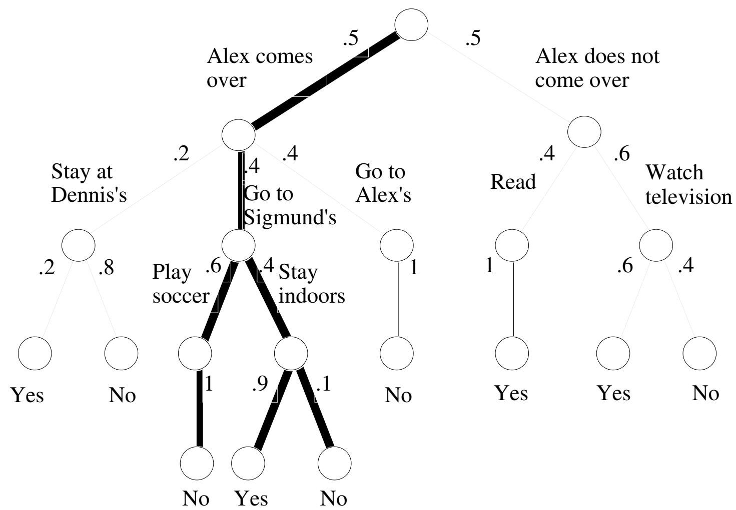


Figure 12 The Moivrean event that Dennis goes to Sigmund's house.

Moivrean events are actually Humean events of a special kind. They are Humean events consisting of chains that begin in the initial situation of the probability tree under consideration and end in a terminal situation. From the point of view of a broader tree, they are situated, but they are all situated in the same context. It is this constancy of situation that makes them ill-suited to play the role of causes.

10 The Character of Statistical Investigation

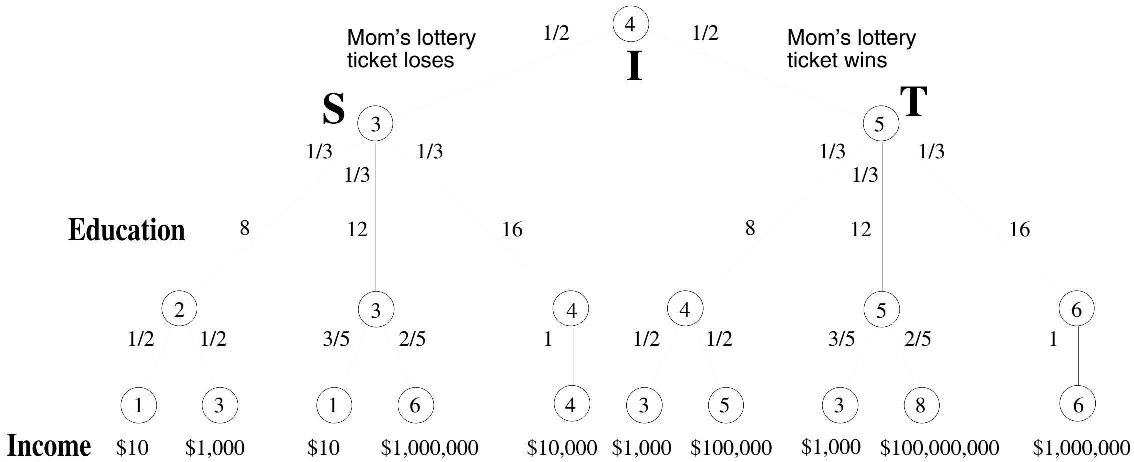
We do not follow the unfolding of events step by step as Nature does. We do not observe Nature's tree. Yet we would like to make guesses about Nature's tree based on what we do observe. It is important, therefore, to identify features of Nature's tree that are closely enough related to what we do see that we can detect them.

This is where variables play an important role. Variables and Moivrean events are things we can observe from outside Nature's tree, and hence things we can use to describe that tree without being inside it. For example, we can pick out steps in Nature's tree—the real contingent causes of some effect—by mentioning some variable that changes in expected value in those steps. We can use the variable "total amount smoked" to pick out steps where a person smokes (or experiences something that increases their likelihood of smoking). This total is merely a number, not a cause of ill-health, but it points to some of the causes quite effectively.

Many standard tools of statistical investigation can be thought of in this way. The coefficient of linear regression of a variable Y on a variable X in situation S, for example, turns out to be a weighted average of the ratio

$$\frac{\text{change in expected value of Y}}{\text{change in expected value of X}}$$

over all the steps where X changes in expected value. Each step is weighted by the square the change in the expected value of X and by the probability (in situation S) that the step will be taken. If the numbers being averaged are quite disparate, then the average is not very interesting, and we will say that the regression coefficient is causally meaningless; it has no simple causal interpretation. But if the numbers being averaged are quite similar, then we will feel that the average does have causal meaning. The most interesting case is where the numbers are all equal. (Figure 13 gives an example.) In this case, the regression coefficient will be the same in every situation, and we will say that there is a stable causal relation between X and Y. This relation deserves a name; I have proposed that we call X a *linear sign* of Y. The concept of linear sign is much more general than the concept of linear causality usually considered by statisticians (e.g., Freedman 1991). But we would be ill-advised to say that X is a cause of Y when X is a linear sign of Y, not only for the general reasons discussed in Section 7, but also because linear sign is only one of many interesting causal relations among variables (Shafer 1996: Chapters 8-10).



The expected value of Log_{10}(Income) is shown in each situation.

Figure 13 In this society, a person's expected income goes up when the person gets more education. This is reflected by the fact that whenever the expected value of education changes, the expected value of the logarithm of eventual income changes by one-fourth as much. Consequently, the regression coefficient of Log_{10}(Income) on education is one-fourth in every situation. The linear regressions are

$$\text{Log}_{10}(\text{Income}) = 1 + \frac{1}{4}\ \text{Education}$$

in the initial situation I,

$$\text{Log}_{10}(\text{Income}) = 0 + \frac{1}{4}\ \text{Education}$$

in situation S, and

$$\text{Log}_{10}(\text{Income}) = 2 + \frac{1}{4}\ \text{Education}$$

in situation T.

Acknowledgement

This paper was inspired by a conversation with Nancy Cartwright. It has benefited, I hope from the good sense of the comments she made in that conversation.

References

Cartwright, Nancy (1989). *Nature's Capacities and Their Measurement.* New York: Oxford University Press.

Dawid, A.P. (1984). Statistical theory: the prequential approach (with discussion). *Journal of the Royal Statistical Society, Series A* **147** 278-292.

Dempster, A. P. (1988). Employment Discrimination and Statistical Science (with discussion). *Statistical Science* **2** 141-195.

Finkelstein, Michael O., and Bruce Levin (1990). *Statistics for Lawyers*. New York: Springer Verlag.

Freedman, David A. (1991). Statistical models and shoe leather (with discussion). *Sociological Methodology* **21** 291-358.

Mackie, J. L. (1974). *The Cement of the Universe.* Oxford: Oxford University Press.

Raiffa, Howard (1968). *Decision Analysis.* Reading, Massachusetts: Addison-Wesley.

Shafer, Glenn (1996). *The Art of Causal Conjecture*. Cambridge, Massachusetts: The MIT Press.