Probabilistic Causation: Some Foundational Issues

- Administrative: (i) please post for this week, (ii) I will lead discussion this week, (iii) paper topics (try to think of one soon).
- Probabilistic Causation: The Basic Idea
  - Causes as probability raisers of effects
  - The symmetry problem
  - The transitivity problem
  - Reflexivity and distinctness
- Spurious Correlations & background contexts
  - Correlation $\not\Rightarrow$ Causation
  - Refinements of basic idea
  - "Causal background contexts"
  - Context Unanimity
  - Reduction and Circularity issues

Probabilistic Causation: The Basic Idea(s)

- The basic idea behind probabilistic theories of causation is that causes C raise the probability of their effects E (here, we remain neutral as to what C and E are, aside from them being "events" in a probability space).

  (PR) $\qquad C$ causes $E \iff \Pr(E \mid C) > \Pr(E \mid \neg C)$.

- C causes E iff C and E are correlated under a suitable probability function Pr. Presumably, "suitable" Pr's will be *objective* probabilities. But, which *particular* Pr's will be appropriate? This is non-trivial...
- Four facts about correlation (each leading to problems):
  - Correlation is reflexive (X causes X?)
  - Correlation is symmetric (X causes $Y \Rightarrow Y$ causes X?)
  - Correlation is *in*transitive (X causes Y & Y causes $Z \not\Rightarrow X$ causes Z?)
  - There are many non-equivalent measures of *degree* of correlation (I.J. Good proposes one, Ellery Eells proposes another)

The Symmetry Problem I

- Correlation is symmetric. But, many think that causation is *a*symmetric (*e.g.*, asymmetric in time). Some (like Suppes and Eells) add an explicit temporal precedence requirement to (PR). This has several problems:
  - This rules out the possibility of backwards-in-time causation *a priori* (many believe this is only a *contingent* fact about causation).
  - This rules out the possibility of a (non-circular) causal theory of temporal order (which has seemed attractive to some philosophers).
  - It is not so clear what it means to say that one property or event type occurs before another ("smoking precedes lung cancer"?).
- Reichenbach takes another (more ambitious) tack. He tries to explain the asymmetry of causation/time, using probabilistic ideas and structures.
- For Reichenbach, the key to temporal/causal asymmetry is in the prevalence of certain kinds of *probabilistic* "fork" structures...

The Symmetry Problem II

- Suppose that factors A and B are positively correlated. And, suppose, moreover, that there is some factor C having the following properties:

  (1) $\qquad \Pr(A \,\&\, B \mid C) = \Pr(A \mid C) \cdot \Pr(B \mid C)$

  (2) $\qquad \Pr(A \,\&\, B \mid \neg C) = \Pr(A \mid \neg C) \cdot \Pr(B \mid \neg C)$

  (3) $\qquad \Pr(A \mid C) > \Pr(A \mid \neg C)$

  (4) $\qquad \Pr(B \mid C) > \Pr(B \mid \neg C)$

- Then, ACB is a *conjunctive fork*. (1) and (2) stipulate that C *screens-off* A from B. This typically occurs when C is a *common cause* of A and B.
- Reichenbach's Theorem: conditions (1) – (4) entail that A and B are correlated, so in some sense C *explains* the correlation between A and B.
- If C is earlier (later) than A, B, and there is no C satisfying (1) – (4) occurring later (earlier) than A, B, then ACB is *open to the future (past)*.

- If a past C and a future D *both* satisfy (1) – (4), then $ACBD$ is *closed*.
- Reichenbach's Proposal: the direction from cause to effect (past to future) is the direction in which open forks predominate. In our world, there are many forks open to the future, few or none open to the past.
- Is the asymmetry between forks open to the past and open to the future as pervasive as this Reichenbach seems to presuppose?
  - In quantum mechanics, there are correlated effects that are believed to have no common cause (no HV proofs) that screens them off.
  - Moreover, if ACB forms a conjunctive fork in which C precedes A and B, but C has a deterministic effect D which occurs after A and B, then $ACBD$ will form a closed fork (see Arntzenius's paper).
  - Also, since this proposal provides a global ordering of causes and effects, it seems to rule out *a priori* backward causation.
- More complex attempts to derive the direction of causation from probabilities have been offered. See Papineau's paper for discussion.

The Transitivity Problem

- Correlation is *not* (in general) transitive. Many philosophers think that causation *is* transitive. How can (PR) cope with this?
- Eells and Sober (see Eells' book *Probabilistic Causality*) prove an important theorem about transitivity and probabilistic causality.
- Let $A \perp B \mid C$ assert that C screens-off A and B. And, let $A \overset{+}{\dashrightarrow} B \mid C$ assert that A and B are positively correlated, conditional on C (and $\neg C$).
- Consider a "chain" of events ABC, such that:
  - $A \overset{+}{\dashrightarrow} B \mid \top$ (*i.e.*, A and B are *unconditionally* correlated)
  - $B \overset{+}{\dashrightarrow} C \mid A$
  - $A \perp C \mid B$
- Then, $A \overset{+}{\dashrightarrow} C \mid \top$. In the standard theories of probabilistic causation (like Eells'), this means that *singly-connected (Markov) probabilistic causal chains are transitive*. Markovian chains are *ubiquitous* in Bayesian nets!

Reflexivity and Distinctness

- Probabilistic correlation is reflexive. As such, naive application of the basic (PR) idea would lead to reflexive causation. This would be wrong.
- We have seen similar issues arise with counterfactual-based theories. Of course, $\neg E \mathrel{\Box\!\!\rightarrow} \neg E$, so counterfactual dependence is also reflexive.
- The standard move at this point is to require that the relata of the causal relation be *logically distinct*.
- Interestingly, many accounts of probabilistic causation do *not* make an explicit assumption of this kind.
- The theories of Suppes and Eells get irreflexivity for free, as a consequence of the temporal precedence requirement.
- Theories which try to get asymmetry without piggy-backing on a temporal ordering will need to add an explicit distinctness requirement to avoid reflexive causation (on a par with counterfactual theories).

Spurious Correlation I

- As we discussed last week, correlation (*simpliciter*) does not imply causation. We need to have correlation on an "appropriate" Pr.
- As a first attempt to handle the problem of spurious correlation, we could try an account of the following kind:

(NSO) Factor A occurring at time t, is a cause of the later factor B iff:

1. $\Pr(B \mid A) > \Pr(B \mid \neg A)$
2. There is no factor C, occurring earlier than or simultaneously with A such that $A \perp B \mid C$ (*i.e.*, no prior screener-off of A, B).

- (2) is intended to eliminate cases where spurious correlations give rise to factors that raise the probability of other factors without causing them.
- But, spurious correlations can also lead to cases where a cause does *not* raise the probability of its effect. So genuine causes need not satisfy (1).
- The general problem here is that Pr must be "causally appropriate" ...

Spurious Correlation II

- (1) & (2) are usually replaced with the requirement that causes raise the probability of their effects in "causal background contexts":

(PRK) A causes B iff $\Pr(B \mid A \,\&\, K) > \Pr(B \mid \neg A \,\&\, K)$ for every causal background context K.

- A causal background context is a conjunction of factors (that is, we can think of $K = K_1 \,\&\, K_2 \,\&\, \cdots \,\&\, K_n$). When K is conditioned on, its factors are said to be "held fixed" (in a *non*-interventionist sense!).
- The (PRK)-theorist owes us an account of which factors should be "held fixed" (*i.e.*, which factors will make-up the "causal background" K).
- The typical accounts sat that in evaluating the causal relevance of A for B, we should hold fixed other causes of B, either positively or negatively.
- But, we should *not* hold fixed causes C of B which are *intermediaries* between A and B (especially, if ACB is a Markov chain).

- Thus, let us call the set of all factors that are causes of B, but are not caused by A, the set of *independent causes* of B.
- A causal background context K will be a *maximal conjunction* (wrt the independent causes of B) $K_1 \,\&\, \cdots \,\&\, K_n$, each of whose conjuncts is either an independent cause of B, or the negation of an independent cause of B.
- Note that the specification of factors that need to be held fixed appeals to causal relations. This appears to rob the theory of its status as a reductive analysis of causation. But, this is not so clear (see below).
- We could generalize (PRK) to allow for 'negative causes', which lower the probability of their 'effect' in all K's, and for 'mixed' causes, which affect the probability of their 'effects' in different ways in different K's.
- Also, we could (as Hitchcock, Woodward, and Pearl do) generalize the theory to apply to non-dichotomous (even continuous) random variables.
- Counterfactual Approach: A-at-t causes B-at-t' iff (i) $\neg$ (A-at-t) $\Box\!\!\rightarrow$ $\Pr(B\text{-at-}t') = b'$, (ii) A-at-t $\Box\!\!\rightarrow$ $\Pr(B\text{-at-}t') = b$, and (iii) $b' \ll b$.

Context Unanimity

- Some theorists (for instance, Eells) require that A and B be correlated in *all* causal background contexts. This is called *context unanimity*.
- Other theorists (*e.g.*, Skyrms, Sober) require only *pareto*-unanimity (*positive* correlation in *some* K's, and *negative* correlation in *none*).
- Critics of these requirements (*e.g.*, Dupré) ask us to

  > Suppose that there is a gene that has the following unusual effect: those that possess the gene have their chances of contracting lung cancer lowered when they smoke. This gene is very rare, let us imagine – indeed, it need not exist at all in the human population, so long as humans have some non-zero probability of possessing this gene (perhaps as a result of a very improbable mutation).

- In this scenario, there would be K's (those that hold fixed the presence of the gene) in which smoking *lowers* the probability of lung cancer: thus smoking would *not* be a cause of lung cancer according these accounts.
- Nonetheless, it seems unlikely that the mere possibility of such a gene would lead us to abandon the claim that smoking causes lung cancer.

Issues of Reduction and Circularity

- As Hitchcock very nicely explains:

  > In order to determine whether a probabilistic reduction of causation is possible, the central issue is not whether the word 'cause' appears in both the *analysandum* and the *analysans*; rather, the key question should be whether, given an assignment of probabilities to a set of factors, there is a unique set of causal relations among those factors compatible with the probability assignment and the theory in question.

- Following Hitchcock, suppose that a set of factors, and a system of causal relations among those factors is given: call this the *causal structure* CS. And, let T be a theory connecting causal relations among factors with probabilistic relations among factors (*e.g.*, (PRK)).
- Then the causal structure CS will be *probabilistically identifiable relative to* T, if for every (this can be weakened) assignment of probabilities to the factors in CS that is compatible with CS and T, CS is the *unique* causal structure compatible with T and those probabilities.

- Intuitively, T allows you to infer that the causal structure is in fact CS, given the probability relations between factors.
- Given a "reductive" probabilistic theory of causation T, there are many identifiability properties it might have, for instance:
  1. All causal structures are probabilistically identifiable relative to T.
  2. All causal structures having some interesting property are probabilistically identifiable relative to T.
  3. Any causal structure can be embedded in a causal structure that is probabilistically identifiable relative to T.
  4. The actual causal structure of the world (assuming there is such a thing) is probabilistically identifiable relative to T.
- Which (if any) of these is required for T to count as a successful "reduction" of causation to probability?
- One of the problems with traditional philosophical theories T of p-causation, is that it is unclear which causal structures are identifiable, relative to T. This is an advantage of the more recent "network" theories.

Illustrative Examples

- (NSO) is able to block the Berkeley graduate school spurious correlation. F = having placed an "F" on your application's gender box in Fall 1971, R = being rejected from Berkeley graduate school in Spring 1972. We have $\Pr(R \mid F) > \Pr(R \mid \neg F)$. But, in each department, this correlation disappears. So, letting K_i = having applied in November to department i, we have $\Pr(R \mid F \,\&\, K_i) = \Pr(R \mid \neg F \,\&\, K_i)$, for each i.
- But, (NSO) does not give the correct answer in the following case. Suppose that smoking is highly correlated with exercise: those who smoke are much more likely to exercise as well. Smoking is a cause of heart disease, but suppose that exercise is an even stronger preventative of heart disease. Then it may be that smokers are, over all, *less* likely to suffer from heart disease than non-smokers. Letting A represent smoking, C exercise, and B heart disease, $\Pr(B \mid A) < \Pr(B \mid \neg A)$. But, if we conditionalize on whether one exercises or not, this inequality is reversed: $P(B \mid A \,\&\, C) > P(B \mid \neg A \,\&\, C)$, and $P(B \mid A \,\&\, \neg C) > P(B \mid \neg A \,\&\, \neg C)$.
- Let A and B be smoking and lung cancer. Suppose C is a causal intermediary, say the presence carcinogens in the lungs. If A causes B *via* C, then C may screen A off from B: given the presence (absence) of carcinogens in the lungs, the probability of lung cancer may not be affected by whether those carcinogens got there by smoking. This is why we do not hold fixed causes of B *that are themselves caused by* A.