Essay Review:

Cause and Chance: Causation in an Indeterministic World,

Phil Dowe and Paul Noordhof, eds., Routledge, 2004

Clark Glymour

Carnegie Mellon University

&

Florida Institute for Human and Machine Cognition

This well-edited volume contains essays of considerable ingenuity on a difficult subject, the analysis of "x causes y" in indeterministic systems. The volume also provides occasion for reflection on the peculiar state and projects of contemporary analytic metaphysics, and I will begin with that.

For most of the contributions to this volume, the project is this: Fill out "Event X is a cause of event Y if and only if……" where the dots on the right are to be filled in by a claims formulated in terms using any of (1) descriptions of possible worlds and their relations; (2) a special predicate, "is a law;" (3) "chances;" and (4) anything else one thinks one needs. The form of analysis is roughly the same as that sought in the *Meno*, and the methodology is likewise Socratic—proposals, examples, counterexamples, more proposals. The norms of the enterprise seem to be as follows (i) a proposal is defeated if someone can *imagine* a circumstance in which it would be false, or perhaps if one can imagine such a circumstance that is not obviously inconsistent with physical laws; (ii) approximately correct solutions, those which cover most but not all cases, are of no value unless they can be modified to cover all cases; (iii) no account is required of how the relations in the right hand side of a proposed analysis could be known or reliably estimated without knowledge of the left hand side. The project is conceptual *analysis*. It means to exclude hypotheses not of biconditional form relating "cause" and other notions. Implicit definition by axiomatic characterization is not allowed. Further, the concern is with logical form rather than with the patterns and constraints revealed in empirical studies of the psychology of human judgement. I believe the restriction to biconditional forms, to analysis, is delusional, the exclusion of psychology is regrettable and the ingenuity shown in many of these essays would profit from the clarity and easy generation of examples that a bit of computer programming would provide.

…….

The trinitarian *deus* of the metaphysics of causality consists of Hans Reichenbach, Patrick Suppes and the late (and deeply missed) David Lewis. Reichenbach is represented here chiefly through references to his late (and also deeply missed) student, Wesley Salmon, for the idea that causation involves a continuous process through space-time in which some property or material part (in Reichenbach's terminology, a "mark") is preserved or passed. Reichenbach attempted to use his theory to define a direction of causation without appeal to a time ordering, but failed for want of a neutral distinction between putting on a mark and taking off a mark. Followers in his tradition have abandoned that ambition and assume a time order. Essays by Phil Dowe and by Douglas Ehring in this volume are in this tradition.

Suppes' account of *event X causes event Y* is this: X and Y are time indexed with X prior to Y, both occur, and the probability of Y conditional on X and all events simultaneous with or prior to the time of X is greater than the probability of Y conditional on the absence of X (at the time of X) and all events simultaneous with or prior to the time of X. Causes, when they occur, raise specific conditional probabilities compared to when they do not occur. Suppes' did not specify an interpretation of probability: however scientists estimate it will do. So generously understood, his theory accords in some essentials with causal analysis in economic time series, but has various problems, some of which are addressed in Igal Kvart's essay in this volume.

Using a metaphysics postulating possible worlds and a proximity relation among possible worlds, Lewis provided a semantics for a logic of counterfactuals. That logic and its semantics in turn provided an interpretation of the counterfactuals in his proposed analysis of causal relations. Lewis' (first) account is this: X causes Y provided they both occur and there is a chain of events X, X1,…, Xn, Y between X and Y such that each successive event in the chain would not have occurred if its immediate predecessor had not occurred. To apply his theory to cases, Lewis required decidable conditions for a proximity ordering of possible worlds, which he provided by working backwards; that is, he specified whatever conditions on proximity relations would result in intuitively true counterfactual claims coming out true according to his theory. One may wonder about the explanatory value of the result of such a procedure, but there it was. Later, he generalized the account to indeterministic worlds, requiring that causes increase the chance of their effects, engendering the objection that some causes lower the chance of their effects. Essays by Stephen Barker, Helen Beebee,, Murali Ramchandran and Paul Noordhof are in this tradition.

Some essays are outliers. Dorothy Edgington does not attempt an analysis but something closer to an implicit definition of "cause." Christopher Hitchcock's essay draws on Judea Pearl's (2000) proposed analysis of actual (or singular or token) causation for deterministic systems. Michael Tooley proposes probabilistic constraints in terms of which he proposes "cause" can be defined.

……

Dorothy Edgington says her aim is to explain why counterfactuals are useful. Her answer is that we use them in hypothetical reasoning both to infer occurrences and to infer non-occurrences, although her illustrations leave out the most obvious use, in planning. Along the way, she argues that however the proximity relation between possible worlds is determined, there are true counterfactuals inconsistent with Lewis' theory of counterfactuals. She proposes instead to modify Lewis' account of counterfactuals as follows: Index events and their absence by time. Suppose neither *A* nor *C* is true. "consider those *A*-worlds which (a) depart from the actual world shortly before the time of ~*A*; .(b) thereafter obey the actual laws of nature; and (c) share with the actual world subsequent particular facts which are causally independent of ~*A* up to the time of the consequent. A counterfactual *A* => *C* is true iff *C* is true in all such worlds." (p. 21). Noordhof disputes that this theory deals adequately with indeterministic cases. I offer no judgement.

Phil Dowe offers two examples of causes that lower the probability of an effect: he shoots a gunman who would shoot and kill you if Dowe had not shot him, but Dowes' bullet goes on to kill you, and an atom which has two decay paths to a decay product but follows the path in which the decay product is less probable. His positive proposal is that in all cases in which a cause lowers the chance, there are two potential "paths" from the cause to the effect, only one of which is actual, and the chance of the effect on the actual path is lower than the overall chance of the effect given the cause. In effect, he excludes worlds in which an effect E can occur spontaneously, uncaused, with some chance and can also be caused by C with a lower chance, and C prevents a spontaneous occurrence of E. Depending on what the "chance of E given C" means (e.g., whether the chance of E given C is or is not dependent on the occurrence of independent events that are also potential causes of C) he may also exclude cases in which some occurring event D could cause E with some chance, while in fact C causes E, but with a smaller chance. Arrangements of the latter kind are in fact central to a psychological account of human causal judgement for which there is considerable evidence (Cheng, 1997), but psychology doesn't get much play in this literature.

As with Dowe's essay, a principal issue of several of the essays in this volume is the claim that some causes lower the chance of their effects, and how the various theories, or variants of them, may be adapted to that fact, if it be one. None of the essays explain what chances are, but Michael Tooley's essay at least directs the reader to a discussion by David Mellor. We are never told how to estimate chances, or the still more obscure and critical conditional chances, from anything, and, again as with Dowe's essay, the lacunae matter.

Helen Beebee rejects Dowe's analysis, and proposes that we should call chance lowering events "hindrances" rather than causes. Why the terminology matters—why "hindrances" rather than "causes that lower the chances of their effects"--is unclear to me. Christopher Hitchcock gives essentially the same analysis as Dowe, differing only in this: whereas Dowe requires that each pathway be a (possible) material process, Hitchcock does not, requiring only that the relevant counterfactuals be true. Very roughly, Hitchcock allows unmediated action at a distance, and Dowe does not. Again, psychological facts would seem to be relevant (do young children who can work TV remotes and wall switches just fine think there is an invisible material process at work?) but none are considered.

Following, like Dowe, the Reichenbach tradition, Douglas Ehring explores a version of what he calls "transference theory": "c causes e just in case there is a transfer of energy momentum from the c-object /location'temporal part to the e-object/location temporal part, with causes and effects consisting of the manifestations of energy/momentum [and]…the quantity 'transferred' persists, at least in part, through the transfer" (p. 72). This is a good theory for football, but one wonders how it works with other field theories: I think the mass of the sun causes the motions of planets, but it doesn't *transfer* any energy or momentum to them, the energy-momentum tensor of empty space is zero. Ehring's solution is that "Causes are connected to their direct effects by way of persisting or partially persisting tropes." (p. 73) "Trope" is a term of art here, meaning a particular that is a property, as the brown of my desk. Whether this addresses the problems of field theories of physics, or not, of itself it seems not to distinguish cause and effect from a sequence of effects such as my shadow as it follows me on a sunny afternoon—the very sorts of "pseudo-causal" processes that material transfer theories were designed to exclude.

Michael Tooley contributes the longest and most complex essay of the volume, full of interesting arguments and observations, with some of which I agree, but to which I cannot even do injustice in this review. So I pass, and leave it to readers. Stephen Barker is not concerned with probability raising or lowering, but with something else—whether one occurring event is a cause (or hindrance in Beebee's terms) of another occurring event. He gives a very lucid account of a number of standard problems about causation—preemption, trumping, overdetermination, transitivity and intransitivity—and proposes a theory of token causation. The "simple" version is the following: c causes e if and only if (i) both occur, and (ii) there are possible events f, such that if f should not occur then c and e would occur, and (iii) if f should not occur then if c should not occur then e would not occur, and (iv) there is no event g that does not in fact occur such that if f should not occur then g would occur and (v) if f should not occur and g should not occur then e would not occur. He illustrates his account with "neuron" diagrams—introduced by Lewis and also used by several other authors here--in which nodes represent variables with two values accordingly as an associated event does or does not occur, and a directed edge indicates that the value of the head node is a function of the value of the tail node (Lewis restricted the functions to "or" functions of all positive causes conjoined with "not" functions of any inhibitory causes, but the graphical representation can of course be used with a broader class of deterministic functions as well as with stochastic causes, as in graphical causal models (Spirtes, et al., 2000; Pearl, 2000)). I do not understand how the five conditions above are realized in neuron diagrams—Barker's treatment of examples sometimes makes up extra nodes as he goes along, which, absent an sunambiguous principle for doing so, seems a cheat, and in cases of overdetermination—four soldiers fire a bullet each of which strikes a man at the same moment and each of which suffices to kill him—requirement (ii) above seems to fail unless events can be disjunctive. This is a case in which an explicit, programmable algorithm would be of help in understanding the proposal.

Igal. Kvart's essay is a genuinely brilliant attempt to deal with inadequacies in Suppes ' account, in particular with the fact that an event C may increase or decrease the chance of E, but that chance may be reversed by one or more other events, not themselves effects of C, occurring after C but prior to E. The obvious problems of this kind are cases in which an intermediate event I interacts with the conditions produced by C to produce E. The essay requires some patience, but repays it. I foresee some difficulties, most obviously how to deal with overdetermination, which Kvart does not address.

Noordhof and Ramachandrian's essays are related attempts to adapt Lewis' theory to complex counterexamples. Noordhof's proposal for deterministic systems goes like this: E1 is relevant to e2 with respect to a set Z of possible events if and only if (i) if e1 and the events in Z were not to occur, e2 would not occur; and (ii) if e1 were to occur and all the events in Z not to occur, e2 would occur. E1 is a cause of e2 if and only if e1 and e2 occur, there is a set Z relative to which e1 is relevant to e2, and all of the events relevant to e2 with respect to Z are those with their actual values, whether they actually occur or actually do not occur.

The proposal is ingenious and accounts for a wide variety of deterministic cases. Noordhof offers a much more complex generalization for indeterministic cases. As with Barker, his deterministic theory appears to require that extra events be interpolated to avoid counterexamples. For example, if e = (~y & x) v w, where the causes are on the right hand side of the equalities, then in the case in which e, x, y and w all occur, x counts as a cause of e, which is implausible. (Let Z = {y occurs, w occurs}). If, however, an event b is inserted between y and e, so that b = ~y and e = (b & x) v w, then x does not count as a cause. This sort of solution requires some principle for inserting additional causes. He proposes (private communication): *If not-Y is part of what is required for E, then there is an event Ei such that Ei is a cause of E and Y inhibits Ei.* "Required for" needs to be cleaned up a bit, perhaps to mean "necessary part of a sufficient condition for E," since not-y is not strictly required for the occurrence of e in the example, but however this is done it seems to necessitate that the effects of negative causes have a dense ordering, since on substituting Ei for E in the italicized sentence, we find that an event Ei', inhibited by Y, must be postulated between Y and Ei. Once again, an algorithm and running program would help us understand the properties of the theory.

Noordhof's theory and Ramachandrian's, like all of those building on Lewis' theory for indeterministic systems, also need to impose restrictions relating chance and causality. If event A raises the chance of B, it might be that A does not cause B but there is a common cause (or a disjunction of several common causes) C of A and of B, and conditional on A the chance of C increases, and conditional on C the chance of B increases. Alternatively, it might be assumed that the chance of an earlier event is zero or one conditional on a later event, but that requires the further assumption that there are no simultaneous causes. If, instead, chances are somehow restricted so that effects do not increase the chances of their causes, then appeal to "chances" rather than to probabilities seems simply a way of sneaking in causal notions on the wrong side of the analysis.

…..

Two things are striking. First, most of these discussions, while claiming to be conceptual analyses, are in fact quite elaborate, substantive theories that cannot plausibly be given biconditional form. The content of most of the ancillary hypotheses is suppressed, but necessary in avoiding counterexamples, and in many cases their truth is not obvious. Second, we would learn much more—or know that we had learned much less--if these theories were written as explicit algorithms and programmed to address neuron diagrams and, for the indeterministic cases that are the cynosure of this volume, graphical causal models as developed in Spirtes, et al. (2000) and Pearl (2000). That would, for example, require metaphysicians to be clearer about chances.[1] The counterfactuals in neuron diagrams and graphical causal models correspond to interventions in the causal system; programs for computing the probabilities that result from such interventions are available as freeware.[2] I suspect David Lewis would have liked resourceful metaphysics all the better were it programmed.[3]

References

P. Cheng, (1997) From Covariation to Causation: A Causal Power Theory. *Psychological Review*, 104, 367- 405.

J. Pearl (2000) *Causality*. New York, Oxford University Press.

[1] For a simple example, see http://www.phil.cmu.edu/projects/actual_causation

[2] At http://www.phil.cmu.edu/projects/tetrad

[3] I thank Paul Noordhof and Douglas Ehring for helpful communications.

P. Spirtes, C. Glymour and R. Scheines, *Causation, Prediction and Search*, 2nd edition, Cambridge, MA., MIT Press.