AI Magazine Volume 19 Number 1 (1998) (© AAAI)

Review of *The Art of Causal Conjecture*

David Hand

In 1992, Glenn Shafer completed a draft of three chapters for an intended book on join-tree computation methods for combining evidence and other aspects of computation in expert systems. However, he found his attention increasingly distracted by the possibilities provided by probability trees for understanding probability and causality—so much so, in fact, that instead of finishing the first book, he wrote a different one on this second topic. It is this second book that is the subject of this review, and it is easy to see why the power and breadth of the ideas seduced Shafer from his original purpose. (Do not despair, however, those three draft chapters have now also appeared, although without the others that were originally intended to accompany them, in *Probabilistic Expert Systems,* Society for Industrial and Applied Mathematics, 1996).

The author describes *The Art of Causal Conjecture* as presenting "a new mathematical and philosophical foundation for probability" (p. xv) and seeking to show how "this foundation can help us understand causality" (p. xv). This is a large claim. Does the book live up to it?

First, of course, there is nothing new under the sun; so, the question arises about the extent to which the presented foundation is new. The central notion of the book is the *probability tree*. These trees have *nodes* that represent the points at which events with more than one outcome can occur and *branches* that represent each possible outcome, with attached numbers signifying the probability of each outcome. This structure is old—Shafer shows that it was familiar to Huygens in the seventeenth century—and indeed it is one to which the reader might have been exposed when first learning about probability. Shafer argues, and the length and depth of the book convinces, that this basic structure is a powerful way of thinking about probability and a way that naturally links probability and causality. He distinguishes the traditional events of probability theory, which he calls *Moivrean events* and which correspond to sets of leaf nodes (or, equivalently, of paths from the root to a leaf), from *Humean events,* which are steps, or chains of steps down through the tree. The distinction, which is clear when presented in the context of a probability tree but less clear from the sample space perspective, is fundamental to the discussion of causality presented in the book.

The Art of Causal Conjecture, *Glenn Shafer, MIT Press, Cambridge, Massachusetts, 1996, 511 pp., ISBN 0-262-19368-X.*

He takes the philosophical orientation that probability is subjective—not in the sense that it depends on someone's state of mind but, rather, that it depends on the perspective from which one sees the world. Probabilities thus describe the position of the observer relative to the world as well as describe the world. Objective probabilities fit nicely into this interpretation because these probabilities describe the view of the world from the position of someone who can see all there is to be seen.

From these simple beginnings, the theory is developed to include different kinds of independence, tracking events and variables, the theory of types, and martingale trees, building up to, toward the end, principles of causal conjecture, causal models, and causal diagrams. This book probably contains more definitions than any other book I can recall reading; however, despite these definitions, one of the attractive features of the book is its essential simplicity. One might have assumed that a deep understanding of the conventional paradigm of the measure-theoretic approach to probability was necessary for a good understanding of the book, but this is not the case. Appendixes do present an overview of the conventional paradigm, and there are links throughout the book between this and the new approach presented here, but the book can be read without a deep grasp of the conventional paradigm. It will be interesting to see how future generations come to regard the ideas in this book. The author might have been especially astute in publishing it in an AI series, so that perhaps more susceptible minds, less encumbered by the paraphernalia of the traditional views, are exposed to the ideas. This, however, is not the reason he gives for publishing it as part of an AI series. The book could have appeared in various subject-matter series, ranging from statistics to computer science, but Shafer chose an AI series because AI, he believes, is currently in the best position to put the book's ideas to use. (Note the word *use*; it implies that he intends the book not simply to be a purely theoretical treatise on a novel way of looking at probability.) He argues that the debate about causal reasoning has been especially intense in the AI and expert system community, so ideas that combine issues of probability and causality fit most suitably into this milieu. Speaking as a statistician, I think it is revealing, true, and something of an indictment of the discipline of statistics that the author can say that causal ideas have not been incorporated into the basic theory of statistics.

The theory described in the body of

the book (that is, distinct from the traditional material relegated to the appendixes) has not been published before. This approach might have the advantage that it will gain more rapid recognition for the achievement represented by the book because the accessible style means that the ideas can be understood by even the relative inexpert; it is not necessary for the material to go through the standard stages of being read by an expert, digested, and translated into simpler terms for the less knowledgeable. The author has already taken these steps. That being said, however, the simplicity of the book, in the sense that it does not depend on years of detailed mathematical study to read and understand, should not detract from its depth. To pick just one example, the notion of weak independence in probability trees leads to a nice way of making Reichenbach's notions of common causes precise.

To me, this book also seems to provide a superb illustration of the merits of going back and revisiting the writings of the greats of the past; for example, "the author first formulated the probability-tree definition of independence and proved Proposition 5.1 in the early 1980s, in the course of trying to understand the seventh definition in Bayes's essay" (p. 133).

This book will amply repay careful study. Given the easy style and attractive presentation, such study should be a great pleasure as well as a thoroughly rewarding exercise.

David Hand is professor of statistics and head of the Statistics Department at the Open University, United Kingdom. He is founding editor and editor in chief of *Statistics and Computing*. He has published over 150 papers and 15 books, including *Artificial Intelligence and Psychiatry, Discrimination and Classification, Artificial Intelligence Frontiers in Statistics, Construction and Assessment of Classification Rules,* and *Statistics in Finance*. His chief methodological research interests are in the foundations of statistics, multivariate statistics, and the interface between computer science and statistics. His main applications interests are medicine, finance, and psychology. His e-mail address is d.j.hand@open.ac.uk.