pragmatic approach, has no answer for them. Whatever his study of the pragmatics of 'should' language accomplishes, it does not explain why you should.

John Morreall
Rochester Institute of Technology

Clark Glymour, Richard Scheines, Peter Spirtes and Kevin Kelly
Discovering Causal Structure: Artificial Intelligence, Philosophy of Science and Statistical Modelling.
San Diego: Academic Press 1987.
Pp. xvii+394. US $39.50.
ISBN 0-12-286961-3.

Glymour et al's book is an interdisciplinary account of the application of the methods of Artificial Intelligence to the problem of extracting causal models from statistical data. This is an enterprise which most undergraduate students in the social sciences (including psychology) are advised against *in principle: correlation is not causation.* However, mainly from the example of their peers, it is a peice of advice they subsequently learn to ignore *in practice.* Glymour et al's book provides (i) a sustained philosophical justification for this practice, and (ii) practical tools for its principled application (the TETRAD program, on a 5.25 inch floppy disc for the IBM-PC and compatibles, comes with the book at no extra cost).

The book is in three parts. The Foreword by Herb Simon locates the work directly in the tradition of employing heuristic techniques in Artificial Intelligence to the problem of scientific discovery. In part I, the basic rationale of the book is outlined and defended. Chapter 1 outlines the general problem of searching for causal models in the social sciences. There are two problems with inferring causal structure from non-experimental data. First, introducing explicit causal models provides for no better predictive success than simple statistical algorithms. Second, the possible causal models of a non-experimental data set grows explosively with the number of variables considered. Glymour et al's Artificial Intelligence solution to these problems introduces the two main themes of the book:

(i) *Explanation*, via the invocation of latent (non-measured) variables (e.g., gravitational forces, electromagnetic fields, etc.), is the primary goal of causal modelling. Therefore, models should be preferred which offer the best explanation of the data. Formal properties of explanation can be represented as mathematical relations between data and model. These relations provide evaluation functions for deriving estimates indicating whether one model provides a better explanation of the data than another.

(ii) The estimates can be used to implement a *heuristic search* of the space of possible elaborations of an initial model. The estimates indicate the best elaborations to explore rather than attempting an impossible exhaustive search of all possible elaborations.

Chapter 2 is a sustained philosophical defence of causal modelling of non-experimental data. The arguments turn on the similarities between the discovery procedures used in the natural sciences and those criticised in the social sciences. Glymour et al argue that the main objections hinge on particular bad practices rather than on any in-principle objection to inferring causes from statistical data. The application of contemporary history and philosophy of science in resolving spurious methodological disagreements in the social sciences is to be applauded. Chapter 3 offers a similar bravura defence of computer-aided discovery.

In part II, the TETRAD program is introduced, primarily in chapters 4 and 5. Causal models are represented as *directed graphs* and a procedure is described for extracting linear statistical models directly from a causal model. The relations between causal model and statistical data which are central to TETRAD's operation involve two sources of constraints, implied by a particular causal model, on the correlation matrix. A model may imply that certain partial correltaions in the matrix should be zero, or that certain *tetrad* equations should apply. In both cases measures of whether the data conform to these constraints can be derived.

Three principles of scientific explanation are identified: *Spearman's principle:* the model should imply constraints that hold in the data; *Thurstone's principle:* the model should not imply constraints which do not hold in the data; *The simplicity principle:* as few causal connections as possible should be invoked. Whether and to what extent the current model satisfies these principles in the data can be assessed from the above measures. The program is interactive and allows users to weight the principles as they see fit.

Starting with an initial or *skeletal* model, elaborations can be effected by adding edges or vertices to the graph representing the model.

Each addition can be assessed to see whether and how far the model is an improvement on the previous model and whether any further additions may effect further improvements, thus effecting a heuristic search through the space of possible causal models. The simplest skeletal model embodies the simplest explanatory hypothesis that all the measured variables are direct causal consequences of one latent variable. However, depending on the prior knowledge of the researcher, more complex initial skeletal models can be employed which include more latent variables.

Chapter 6 provides an impressive account of TETRAD's ability to locate causal models. Chapter 7 follows up with some simulation studies, and Chapter 8 with some particular case studies drawn from the social sciences. Chapter 9 is a brief historical account of heuristic search in Applied Statistics, where Glymour et al trace the antecedents of the TETRAD philosophy in the work of Spearman. Chapter 10 provides the relevant proofs of the various mathematical results on which TETRAD relies.

Part III is the TETRAD manual, which provides a lucid and thorough account of the program. Chapter 12 discusses how to employ TETRAD with other programs such as EQS and LISREL. That TETRAD can not be used independently may be seen as a disadvantage; however, those researchers for whom TETRAD will be of most practical value will already be familiar with these programs.

TETRAD is an important and valuable addition to the store of exploratory statistical techniques available to social scientists. Moreover, the book provides the TETRAD user with a sustained philosophical justification for its employment. In this sense *Discovering Causal Structure* is a unique book. It is exemplary to historians and philosophers of science who may wonder whether their discipline has an applied dimension. It is of direct interest to workers in Artificial Intelligence as an example of a practical application of the methods of heuristic search. Moreover, it may prove of immense practical value to researchers in the social sciences concerned to derive better explanations of their data. Thus, *Discovering Causal Structure* is not only a necessary addition to the library, it should also adorn the bookcases of disparate researchers whose collections had previously been mutually exclusive.

Mike Oaksford
(*Department of Psychology*)
University College of North Wales