Reviewed by Jun Huang, Advanced Computation Laboratory, Imperial Cancer Research Fund, London, UK

Probability and the art of judgement by Richard Jeffrey, Cambridge University Press, 1992, pp 244, £12.95 (p/back) or £37.50 (h/back), ISBN 0-521-39770-7.

For anyone who is interested in decision theory, and like me finds the name of Richard Jeffrey far more familiar than his work, this volume of essays is essential reading. Despite being a collection of pieces, and therefore lacking a common narrative thread to tie the various ideas tightly together, it nevertheless manages to give a comprehensive treatment of the topics with which he is concerned. Broadly speaking, these are all aspects of the use of probability and utility theory to make decisions in the face of uncertain information, but Jeffrey's treatment differs from that of other Bayesian probabilists in two important ways. First, Jeffrey approaches the problem very much from the philosophical rather than technical point of view; he is, after all, a philosopher. Second, he gives considerable space to discussing whether the theories provide a reasonable model of human, rather than purely rational, decision making. As a result, the book is always very readable, and the theories he advances more convincing than those built from purely technical considerations.

In total, the book contains 16 essays, the earliest of which was written in 1956, and the most recent of which dates from 1991. Thus, what the book provides is a historical record of Jeffrey's ideas gathered together in one handy volume and ordered so that similar topics are covered in adjacent essays. In addition, through the new introductory essay, and the postscripts to a couple of the others, recent developments are also acknowledged. As a result, this is more than just a collection of papers, and is instead a coherent whole, albeit one that is more of a palimpsest than a conventional monograph.

In brief, the topics of the individual essays are as follows. The introduction argues that the position established by the whole collection is but a development of Carnap's views on probability, and represents a generalization of logic deductive methods in which the concepts of true and false are replaced by a continuum of states of belief. The second essay then addresses the question of scientific reasoning in the light of this position, and reaches the conclusion that the best decisions will be reached only if hypotheses are not merely accepted or rejected, but are assessed probabilistically—exactly the kind of conclusion one would expect from an advocate of Bayesian techniques. However, to be fair, it should be noted that Jeffrey also admits that there are considerable problems with this view, not least among which is the fact that this is clearly not what happens in practice.

The third essay introduces the important idea of probability kinematics. This is a means of updating probabilistic information in the light of uncertain information, and is a generalization of conditioning, which only permits updating in the light of certain information. Clearly the use of probability kinematics makes if possible to carry out probabilistic updating in a much broader set of situations than conditioning, and Jeffrey uses this fact to argue that probability kinematics provide the "probabilism of the commonplace". Then, in the fourth essay, he uses it to form the basis of an account of scientific method. This account is extended in the fifth essay with a description of how scientific theories can become more probable by explaining previously known facts (the tricky problem of old evidence), and how incomplete and indefinite probabilistic information can be usefully employed. Further work on probability kinematics is contained in the sixth and seventh essays. The first of these discusses the order of revisions, and proposes a solution to the problem of revision in the face of evidence in which, unlike other proposals, the order of two successive revisions does not effect the final probability. The second looks at the connections between probability kinematics and subsequent developments in Bayesian statistics, showing various equivalences between the approaches, and further developing the theme of successive updating.

The eighth essay introduces Jeffrey's third major theme after scientific reasoning and probability kinematics—the study of preferences and their representation. Indeed, with the exception of the eleventh essay, which is a convincing demolition of the frequency view of probability, and the twelfth which discusses explanation, all the remaining papers deal with this topic in a series of bite-sized chunks.

First is a discussion of a preference relation that can be used to construct higher order degrees of preference, such as the subtle notion that one can prefer the option of preferring not to smoke over the option of preferring to smoke, despite preferring to smoke. This is then followed by two closely related essays on the use of utility theory to make interpersonal comparisons, in other words, to provide a sensible means of choosing options that, as far as possible, satisfy everybody's preferences. Clearly, this is not straightforward, since it involves using some model of how to assess the well-being of a whole group from the well being of its individual members and the consequent dip in extremely murky waters. The final four essays are even more closely bound together, all being concerned with various aspects of the model of preferences introduced in the first part of Jeffrey's book *The Logic of Decision*. The first of these four, the thirteenth in the collection, summarizes the model in a few pages, and the second briefly compares it with other proposals, not surprisingly concluding that it is superior to them. The fifteenth essay then provides a detailed look at a system of axioms for the model of preference, and the sixteenth and final essay discusses at length the way in which preferences may be updated.

Reviewed by Simon Parsons, Advanced Computation Laboratory, Imperial Cancer Research Fund, London, UK

Inductive logic programming: techniques and applications by Nada Lavrac and Saso Dzeroski, Ellis Horwood, UK, 1993, pp 293, £39.95, ISBN 0-13-457870-8.

Inductive logic programming (ILP) is the fusion of symbolic machine learning with logic programming. Its aim is to learn logic programs from examples. The use of the logic programming formalism has advantages for machine learning. The formalism is elegant with a strong theoretical base, making it relatively easy to learn logic programs compared to, say, learning C programs. The formalism also makes it easier to incorporate background knowledge in the learning process, an essential to solve many learning problems. Logic programs (based on the first-order predicate calculus) are more expressive than traditional symbolic machine learning formalisms (based on the propositional calculus), such as decision trees. This means that a larger class of learning problems can be dealt with. The disadvantage of using such a powerful representation as ILP is that the search space is generally larger than for propositional learning; this is the reason that ILP has been, until recently, considered infeasible.

To appreciate the need for learning systems of the power of ILP, consider the following simple problem (from Quinlan, but very similar to many drug design problems). Consider a directional network of ten nodes, with each node connected to at most three others. In a traditional (propositional) machine learning formalism, an object must be specified by its values for a fixed set of attributes, and the rules learnt expressed as functions of these attributes. This would mean for our example that 30 (3×10) attributes would be needed, e.g. the attributes A1,B1,C1 would be the nodes to which node 1 is linked. Now consider the problem of expressing the concept "two nodes are linked to each other" using the propositional formalism. It gives a very complicated expression:

$$(A1 = 2 \vee B1 = 2 \vee C1 = 2) \,\&\, (A2 = 1 \vee B2 = 1 \vee C2 = 1) \vee$$
$$(A1 = 3 \vee B1 = 3 \vee C1 = 3) \,\&\, (A2 = 3 \vee B2 = 3 \vee C2 = 3) \vee$$
$$(A1 = 4 \vee B1 = 4 \vee C1 = 4) \,\&\, (A2 = 4 \vee B2 = 4 \vee C2 = 4) \vee$$
etc.

$$(A2 = 3 \vee B2 = 3 \vee C2 = 3) \,\&\, (A3 = 2 \vee B3 = 2 \vee C3 = 2) \vee$$
etc.

Now considering the problem in the ILP formalism, using only the predicate linked-to (X,Y) to express the fact that there is a directed link from X to Y. The concept "two nodes are linked to each other" can now be easily expressed by the PROLOG program: