A Review of
"Probabilistic reasoning in expert systems — theory and algorithms"
by Richard E. Neapolitan

Wiley-Interscience Publication, 1990, £46.90
ISBN 0-471-61840-3

reviewed by:

Simon Parsons
Advanced Computation Laboratory
Imperial Cancer Research Fund
P.O. Box 123,
Lincoln's Inn Fields
London WC2A 3PX

My initial feeling on coming across "Probabilistic reasoning in expert systems" was one of amazement. I was astonished that anyone could think that it was a good idea to produce a book that paralleled Judea Pearl's [1988] seminal work on Bayesian networks to the extent of having an almost identical title, and included a re-explanation of Pearl's theoretical work on probability propagation. On sober reflection I have changed my mind. Indeed, I can not only understand why "Probabilistic reasoning in expert systems" was produced, but also I prefer it to Pearl's book, largely because to me it reads more as summary of an area of research rather than a monograph which seeks to advance a particular point of view. Thus Neapolitan takes a less hard-nosed Bayesian approach to the material that he and Pearl have in common, and covers a several useful topics that are not covered in Pearl's work. However, whether this makes up for the massive price difference between the books is something that I'm not entirely sure about.

Neapolitan starts from the position that because expert systems have to deal with complex domains in which the relationships between cause and effect are inexact, the conclusions that the systems deal with are inherently uncertain and one must provide them with some mechanism for dealing with that uncertainty. This leads to a detailed discussion of classical probability theory, including a thorough exploration of the subjectivist and frequentist positions, and this discussion is one area in which, in my view, Neapolitan scores heavily over Pearl . This is because the latter restricts his introduction to probability theory to the subjectivist position. Whilst this is understandable given that as a Bayesian Pearl is not concerned with frequentist ideas, I feel that to disregard the position gives a somewhat biased view of the technology that he goes on to introduce since it is not by any means restricted to subjective probabilities. Similarly, it would be very useful if Pearl had included material on the principle of maximum entropy as Neapolitan does.

Following the material on probability theory Neapolitan turns to the other major theoretical component of Bayesian network— graph theory. The parts of graph theory that are necessary in order to prove the various technical results in Bayesian networks, that is ideas such as paths, cycles, cliques and triangulation, are introduced in a few rather dense pages. Although this material is self-contained in one chapter, it sits rather awkwardly in the book . When one comes across it, it is not really clear why it is needed since the idea of a belief network hasn't been introduced. Having said that, having the graph theoretic material in a ghetto of its own pays dividends when one gets further on in the book since it makes for convenient reference. All things considered, however, it might have been better to put the graph theory after the chapter it preceeds since the latter introduces belief networks, By making suitable comparisons this chapter also demolishes the idea that it is possible to reason with uncertainty in rule-based systems, exposing the well-documented problems that have dogged attempts to do so. Whilst it is good to

see approaches compared in this way, it is a shame that Neapolitan rather overstates his case by ignoring the possibility of using formalisms like probabilistic logic [Nilsson 1986], [Nilsson 1993] and possibilistic logic [Dubois and Prade 1987]. The first four chapters thus form the background material that the book requires, and with the fifth chapter the real business, the exposition of a scheme for representing and reasoning with uncertain information using Bayesian networks, gets under way.

As one might expect, Chapter Five deals with representation issues, defining a network formally and discussing how the probability elements and the graph theory elements of the method fit together. This discussion extends into the following chapter, which includes an exposition of d-separation and its relation to probability, graph theory, and causality. This explanation is so clear that even the misprinting of one of the figures, with the directed arcs pointing in the wrong direction, fails to make it confusing. The rest of the sixth chapter is taken up with Pearl's schemes for propagating probabilities in singly connected causal networks by local computation, including proofs and the pseudocode of the necessary algorithms, and the seventh extends the results to the general case of multiply connected networks. In this chapter Neapolitan departs from covering Pearl's work in detail, choosing instead to discuss the method proposed by Lauritzen and Spiegelhalter [1988] for local computation based upon join trees. Given the comparative popularity of the latter method and those proposed by Pearl for the same purpose, this is a wise choice, and as Pearl admits in a recent paper [Pearl 1993], is the one indicated by hindsight. These three chapters form the heart of the book, and in my opinion they make it well worth reading. The ideas ar e clearly presented and they cover the most widely used techniques, even, in the case of the Lauritzen and Spiegelhalter method, including an improvement to the standard algorithm. One couldn't really ask for more.

The remaining chapters cover a number of related topics. Chapter Eight is concerned with abduction in both singly and multiply connected probabilistic networks, which in this case is defined as being the problem of determining the most probable combination of all uninstantiated variables in the network when certain observations are made. In other words, to put things into the usual medical context, this is the problem of determining what the most likely explanation of a given set of symptoms is. Chapter Nine then extends the work to cover decision making. This involves the addition of nodes representing decisions to causal networks, thus creating influence diagrams and allowing utility calculations to be made on the results of the probability calculation. Finally, Chapter Ten looks at what the probabilities that previous chapters have dealt with actually mean, and proposes the use of Dirichlet distributions to represent uncertainty in probability estimations.

Thus, overall, my view is that "Probabilistic reasoning in expert systems" is a thoroughly good book, and I would have no hesitation in recommending it to anyone who wanted to read up on the basic mechanisms of probabilistic networks. I do have a couple of complaints about it. Firstly, it should be recognised that the use of Bayesian networks is only one of the possible ways of correctly handling uncertainty in expert systems, and this fact is not really acknowledged by Neapolitan. Secondly I don't like the price of the book. I forget how much Pearl's 'Probabilistic reasoning in intelligent systems' cost, but I don't believe that it was more than £30. That means that Neapolitan's volume retails at more than half as much again, which, given that both are nice harbound editions of roughly the same length[1], makes the latter appear ridiculously over-priced.

References

Dubois, D. and Prade, H. (1987) Necessity measures and the resolution principle, *IEEE Transactions on Systems, Man and Cybernetics*, **17**, 474–478.

[1]Indeed, Pearl's book is the longer of the two, including additional material summarising his work on the relation between logic and probability

Lauritzen, S. L. and Spiegelhalter, D. J. (1988) Local computation with probabilities in graphical structures and their applications to expert systems, *Journal of the Royal Statistical Society B*, **50**, 2.
Nilsson, N. (1993) Probabilistic logic revisited, *Artificial Intelligence*, **59**, 39–42.
Nilsson, N. (1986) Probabilistic logic, *Artificial Intelligence*, **28**, 71–87.
Pearl, J. (1993) Belief networks revisited, *Artificial Intelligence*, **59**, 49–56.
Pearl, J. (1988) *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann, San Mateo.