Review of

The Uncertain Reasoner's Companion: A Mathematical Perspective
by J.B. Paris
Cambridge University Press, 1994.

Ernest Davis, New York University

The attainment of a complete formal theory of plausible reasoning, in the sense that classical logic is a complete formal theory of sound deduction, is a goal of the highest importance to AI and to many other fields. The pursuit of this goal has encountered many obstacles and, incidentally, has given rise to a literature of surprisingly bitter and violent polemic. Nonetheless, there is a large and steadily growing body of knowledge about the philosophical, mathematical, computational, psychological, and empirical aspects of this problem. J.B. Paris' *The Uncertain Reasoner's Companion* is a valuable and important survey of the mathematical analysis of quantitative theories of plausible inference, which consists primarily of theorems of the form, "If you wish your inference system to have the apparently desirable properties A, B, and C, then the system must satisfy X, Y, and Z." The mathematics, of course, does not tell you what properties plausible inference *should* have.

Paris grounds the fundamental problem of plausible inference in the following scenario. You are building an expert system, which is supposed to assign degrees of belief, absolute and conditional, to propositions in a language $\mathcal{L}$. You consult with an expert, who tells you his own evaluations of some of the propositions in $\mathcal{L}$. Let K be this collection of evaluations. Then the expert goes away. How, now, should the expert system assign likelihoods to those propositions in $\mathcal{L}$ that are not in K?

At first glance, this scenario, with its two separate reasoners, one human and one automated, seems awkward. Actually, it is remarkably ingenious. Paris does not trumpet the advantages of this setting, so let me do so:

- It is obviously of practical significance.
- It is theory-neutral as between the different quantitative theories of degree of belief, such as probability, Dempster-Shafer theory, and fuzzy logic.
- It sidesteps the extremely controversial issue of the "proper" interpretation of judgments of likelihood.
- It combines a purely subjective interpretation of K (the likelihoods are what the expert says they are) with a consideration of objective criteria for inference rules extending these assignments in the expert system.
- It vitiates the knee-jerk response of the orthodox Bayesian, "Get the needed priors from the expert (by offering him a series of wagers and badgering him into consistency) and use Bayesian updating. Anything else is provably suboptimal." The actual expert is no longer on the payroll to be asked. Nonetheless, some extensions to K seem more sensible than others.

There is, however, one difficulty, which is the issue of the consistency of K. Most inference techniques can only be applied if K is consistent relative to the relevant theory of likelihood. This means that we are not simply taking what the expert tells us at face value. Rather, we are requiring his judgments to observe certain constraints; these constraints presumably draw their force from a particular semantics or interpretation . . . In short, we end up, in any case, to some degree imposing our own theory of likelihood on the expert. As a practical matter as well, of course, it is not at all unlikely that K will be inconsistent, if it is large and unstructured. The expert systems designer is still obliged to do something with it, and what he does is not covered by any of the techniques here.

Having set up the problem, Paris proceeds to examine the constraints on the likelihood function imposed by probability theory (chapter 3), Dempster-Shafer theory (chapter 4), and fuzzy logic and variants (chapter 5); inference processes, particularly maximum-entropy (chapter 6); desiderata for inference processes (chapter 7); techniques for belief revision, particularly minimal cross-entropy (chapter 8); consequences of further imposing independence constraints (chapter 9); computational complexity of plausible inference (chapter 10); and the application of likelihood estimates to first-order languages (chapters 11 and 12). The survey includes well-known fundamental theorems in this area, such as Cox's and de Finetti's justifications of the axioms of probability, and Shannon's and Boltzmann's justifications of the maximum entropy principle, and numerous new results, many of them due to the author.

Paris' presentation is resolutely mathematical: definition, lemma, theorem. Intuitive motivations for the various desiderata for inference schemes are given briefly but are not argued at length. Discussions of non-mathematical aspects of plausible inference is mostly parenthetical. The mathematics background required does not go beyond calculus and a little logic, except in the chapter on computational complexity. (Curiously, probability theory is hardly used.) The book is exceptionally clear and coherent, though spare, by the standards of AI writing.

It is a slim volume, just over 200 pages, with a narrowly defined scope. Most notably, it omits any discussion of randomness or random selection (except in the context of random Turing machines) or of any of the issues that draw on or connect to that Protean concept, such as statistical inference, PAC learning (both of which assume that the data presented is a random sample), or Kolmogorov complexity.

This book should, and I am sure will, become a standard reference and textbook in the field. Having a single volume that collects this material and presents it clearly in a uniform approach is extremely valuable to the researcher and the student. I feel this particularly strongly, as last semester, as it happens, I taught a course on "Probability and AI" of which about a third dealt with the kind of material in this book. I had to scrounge up articles from here and there, to reconstruct theorems and proofs from general descriptions, and so on. All this would have been vastly easier and better had I had this book in hand.