Foreword

Probabilistic models based on directed acyclic graphs have a long and rich tradition, beginning with work by the geneticist Sewall Wright in the 1920s. Variants have appeared in many fields. Within statistics, such models are known as directed graphical models; within cognitive science and artificial intelligence, such models are known as Bayesian networks. The name honors the Reverend Thomas Bayes (1702−1761), whose rule for updating probabilities in the light of new evidence is the foundation of the approach. The initial development of Bayesian networks in the late 1970s was motivated by the need to model the top-down (semantic) and bottom-up (perceptual) combination of evidence in reading. The capability for bidirectional inferences, combined with a rigorous probabilistic foundation, led to the rapid emergence of Bayesian networks as the method of choice for uncertain reasoning in AI and expert systems, replacing earlier ad-hoc rule-based schemes. Perhaps the most important aspect of Bayesian networks is that they are direct representations of the world, not of reasoning processes. The arrows in the diagrams represent real causal connections and not the flow of information during reasoning (as in rule-based systems or neural networks). Reasoning processes can operate on Bayesian networks by propagating information in any direction. For example, if the sprinkler is on, then the pavement is probably wet (prediction); if someone slips on the pavement, that also provides evidence that it is wet (abduction, or reasoning to a probable cause). On the other hand, if we see that the pavement is wet, that makes it more likely that the sprinkler is on or that it is raining (abduction); but if we then observe that the sprinkler is on, that reduces the likelihood that it is raining. It is the ability to perform this last form of reasoning—called explaining away—that makes Bayesian networks so powerful compared to rule-based systems or neural networks. They are especially useful and important for risk assessment and decision-making.

Although Bayesian networks are now used widely in many disciplines, those responsible for developing (as opposed to using) Bayesian network models typically require highly specialist knowledge of mathematics, probability, statistics, and computing. Part of the reason for this is that, although there have been several excellent books dedicated to Bayesian Networks and related methods, these books tend to be aimed at readers who already have a high level of mathematical sophistication—typically they are books that would be used at graduate or advanced undergraduate level in mathematics, statistics, or computer science. As such they are not accessible to readers who are not already proficient in those subjects. This book is an exciting development because it addresses this problem. While I am sure it would be suitable for undergraduate courses on probability and risk, it should be understandable by any numerate reader interested in risk assessment and decision making. The book provides sufficient motivation and examples (as well as the mathematics and probability where needed from scratch) to enable readers to understand the core principles and power of Bayesian networks. However, the focus is on ensuring that readers can build practical Bayesian network models, rather than understand in depth the underlying algorithms and theory. Indeed readers are provided with a tool that performs the propagation, so they will be able to build their own models to solve real-world risk assessment problems.

Judea Pearl

UCLA Computer Science Department

Los Angeles, California