The Knowledge Engineering Review, Vol. 12:4, 1997, 413–415

Directions in uncertainty reasoning

ERIC NEUFELD

Department of Computer Science, 57 Campus Drive, University of Saskatchewan, Saskatoon. Saskatchewan, Canada, S7N 5AS

'Uncertainty reasoning" refers in a general way to problems discussed by that subset of the AI community interested in representing and reasoning with knowledge that cannot be expressed as certainties. The range of problems discussed runs the gamut from fundamental philosophical inquiry into the nature of uncertainty and how (if at all) it can be measured and modelled, to practical performance issues arising from the (automatic) construction of real-world models and making inferences from such models.

A decade ago, this community devoted significant discussion (see, for example, Cheeseman (1988) and Kanal and Lemmer (1986), and more recently, Kyburg (1994)) to philosophical problems regarding the fundamental nature and representation of uncertain information. For instance, the community examined numeric formalisms other than standard probability for representing the degree of uncertainty (for example, fuzzy logic (Zadeh, 1986), certainty factors (Heckerman, 1989) and belief functions (Shafer, 1986), as well as non-numeric formalisms, both probabilistic (qualitative probabilistic networks (Wellman, 1987)) and nonprobabilistic (endorsements (Cohen, 1983)). More recently, the community has focused on probabilistic and statistical approaches in general, including both the model inference problem and the problem of making inferences from models.

Although probabilistic AI has entered the AI mainstream, the breadth and depth of uncertainty reasoning is unfamiliar to many, and the interested reader should pay attention to developments reported in specialized uncertainty forums. A new forum is the FLAIRS (Florida Artificial Intelligence Research Symposium) Special Track on Uncertain Reasoning, which for the past two years has assembled a representative selection of original and high quality papers reporting a wide range of results (Dankel, 1996, 1997)) from the most philosophical to the most practical.

Snow (1996) addressed several foundational problems relating to the '''lottery paradox'' and the ''grue paradox''. The lottery paradox arises when formalisms treat sentences exceeding some probabilistic threshold as ''practically true''. But for any threshold, one can devise a lottery where, for each ticket, it is practically true that the ticket will not win. However, the conjunction of these sentences is inconsistent with the design of the lottery where a single winner is guaranteed. The ''grue paradox'' is an inductive paradox, which might be summarized as follows. The fact that every emerald we have seen is green might be accepted as evidence for the hypothesis that ''all emeralds are green'', but perhaps it is also evidence for the hypothesis that ''all emeralds are grue'', where an object that is ''grue'' is green now, but will change to blue in the year 2222. Snow presents systems based on ''possibility'' and ''ignorant induction'' to avoid this paradox. Tawfik (1986) discussed how to use I. J. Good's ''surprise index'' to handle the issue of representing surprises in probablistic reasoning. This allows us to distinguish surprise about the sentence ''I won the lottery'', or ''my friend won the lottery'' from ''somebody won the lottery''.

Bayesian networks, also known as belief nets or causal nets, have been particularly noticeable in the AI mainstream. Causal nets compactly represent probability distributions of many variables, and offer certain efficiencies with respect to both inference and model construction. Although a recent textbook (Russell & Norvig, 1995) devotes several chapters to their study, the understanding of this technology is advancing rapidly as it is deployed in real world settings. Although they offer some efficiencies, the complexity of inference can grow exponentially with the number of variables in the net. One solution is multiply sectioned nets, which partitions the domain into more manageable subsets. Xiang (1996) spoke to the problem of verification of the structure of partitioned networks that can be managed by cooperating individual agents within the system. Darwiche & Provan (1996) addressed the efficiency issue in a different way by showing how belief networks can be compiled into arithmetic expressions called Query DAGs. Interestingly, Query DAGs can be generated based on any of the popular exact algorithms, but significantly reduce hardware and software requirements for on-line applications.

Qualitative inference has long been of interest to the uncertainty community, incluing the Bayes net community. An interesting range of probabilistic inferences can be accomplished without numbers, using statistical majorities (Bacchus, 1989), extreme probabilities (Pearl, 1988) and probabilistic influences (Wellman, 1987) (statements that assert that knowing some statement is true increases the probability of another). In particular, qualitative influences defined on the arcs of a belief net have the attractive property of propagation, allowing principled transitive inference. Parsons (1997) noted that despite this, a broad class of networks have the problem of over-abstraction, that is, the kinds of inference are still very limited. He showed that introducing qualitative order-of-magnitude reasoning, that is, specifying that certain probabilities were orders of magnitude greater than others, extended the range of interesting inferences. This has an advantage over formalisms using infinitesimals (such as the epsilon-semantics) that the relations can at least in principle be objectively verified. O'Rorke & Henrion (1996) also considered the use of order of magnitude reasoning for qualitative decision theory, in the representation of information about utilities, but concluded that order of magnitude methods alone will not provide a sufficient basis for decision analysis as they leave most problems undecided, they and are considering integrating other approaches into their system.

A common use of Bayes nets is in diagnosis. A problem with diagnostic systems based on logical formalisms such as the ATMS or Theorist, as well as those based entirely on Bayes nets, is that they are not sensitive to probabilities that change over time. For example, many products have a tendency to exhibit failure almost immediately (due, say, to manufacturing flaws), but otherwise last a long time (the so-called "bathtub" distribution). Trying to design a diagnostic system that accommodates uncertainty of such a general nature creates tractability problems. Tawfik & Neufeld (1997) showed how to use logic formalisms and thresholding to, in essence, constrain the focus of a diagnostic algorithm to more probable outcomes at any particular time point. This hybrid formalism can thus evolve with time.

Although the uncertainty community has been interested in the problem of inferring models from data for some years, this has perhaps entered the AI mainstream relatively recently. Some of this also goes under the heading of Knowledge Discovery and Data Mining. Most of the papers at the 1997 special track addressed this problem. Zhong et al. (1997) described a new way to discover rules in data, generalizing on Mitchell's "version spaces". Whereas traditional version space algorithms depend on complete information to produce compact rule descriptions from the data, Zhong showed a method based on Generalized Distribution Tables that allowed incomplete and noisy information by assigning indifferent priors. Konkanen et al. (1997) discussed a Bayesian approach to constructing finite mixture models from sample data. A finite mixture is a mixture of unknown proportions of several populations that differ in significant ways (Everitt & Hand, 1981). These differences can be in essence attributed to latent variables, and the problem of inferring the unknown proportions gives the experimenter insights into the structure of the hidden variable. The system described is entirely unsupervised. Chu & Xiang (1997) also discussed the problem of learning belief networks over large domains by using parallelism and exploiting a multi-link lookahead strategy. This method resulted not only in a speedup due to parallelism, but increased efficiency for each processor devoted to the task.

It is not difficult to view uncertain reasoning not as a subarea of AI, but as the foundation of AI.

Uncertainty techniques have been applied to many, if not all, other areas of AI, including planning and natural language processing, and corporations are devoting research energies to exploring commercial possibilities. Current plans are to expand the FLAIRS special track to allow discussion of this wider variety of subjects in the uncertainty context.

References

Bacchus F, 1989. "A modest, but semantically well-founded inheritance reasoner" *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pp. 1104–1109.

Cheeseman P, 1988. "An inquiry into computer understanding" in MacLeish M (ed), *Computational Intelligence 4*, pp. 58–66.

Chu T and Xiang Y, 1997. "Parallel learning of belief networks" *Proceedings of the Tenth International Florida Artificial Intelligence Research Symposium*, pp. 192–197.

Cohen PR, 1983. "A theory of heuristic reasoning about uncertainty" *AI Magazine* Summer, pp. 17–24.

Dankel II, DD, 1997. *Proceedings of the Tenth International Florida Artificial Intelligence Research Symposium*, Daytona Beach, May 12–14, pp. 186–215.

Dankel II, DD, 1996. *Proceedings of the Tenth International Florida Artificial Intelligence Research Symposium*, Key West, May 20–22, pp. 285–310.

Darwiche A and Provan G, 1996. "Implementing belief-network inference in real-world systems" *Proceedings of the Ninth International Florida Artificial Intelligence Research Symposium*, pp. 290–294.

Everitt BS and Hand DJ 1981. *Finite Mixture Distributions*. Chapman & Hall.

Heckerman D, 1989. "Probabilistic interpretation for MYCIN's certainty factors" in Bacchus F (ed), *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, pp. 167–196.

Konkanen P, Myllymaki P and Tirri H, 1997. "Experimenting with the Cheeseman-Stutz evidence approximation for predictive modeling and mining" *Proceedings of the Tenth International Florida Artificial Intelligence Research Symposium*, pp. 201–211.

Kyburg HE, Jr, 1994. "Believing on the basis of the evidence" in Neufeld E (ed), *Computational Intelligence 10*, pp. 3–20.

O'Rorke P and Henrion M, 1996. "Qualitative decision analysis: evaluating an order of magnitude calculus (some initial results)" *Proceedings of the Ninth International Florida Artificial Intelligence Research Symposium*, pp. 305–309.

Parsons S, 1997. "On order of magnitude reasoning and qualitative probability" *Proceedings of the Tenth International Florida Artificial Intelligence Research Symposium*, pp. 198–203.

Pearl J, 1988. *Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference*. Morgan Kaufmann.

Russell S and Norvig P, 1995. *Artificial Intelligence: A Modern Approach*. Prentice-Hall.

Shafer G, 1986. "Probability judgement in artificial intelligence" in Kanal LN and Lemmer JF (eds), *Uncertainty in Artificial Intelligence*, pp. 127–136.

Snow P, 1996. Inductive reasoning with conditional probabilities" *Proceedings of the Ninth International Florida Artificial Intelligence Research Symposium*, pp. 300–304.

Tawfik A and Neufeld E, 1997. "Surprises in probabilistic reasoning" *Proceedings of the Tenth International Florida Artificial Intelligence Research Symposium*, pp. 285–289.

Tawfik A and Neufeld E, 1996. "Changing defaults in diagnosis" *Proceedings of the Ninth International Florida Artificial Intelligence Research Symposium*, pp. 212–216.

Kanal L N and Lemmer J F (eds), 1986. *Uncertainty in Artificial Intelligence*. North-Holland.

Wellman MP, 1987. "Qualtitative probabilistic networks for planning under uncertainty" in Kanal LN and Lemmer JF (eds), *Uncertainty in Artificial Intelligence 2*. North-Holland, pp. 198–208.

Xiang Y, 1996. "Distributed structure verification in multiply sectioned Bayes networks" *Proceedings of the Ninth International Florida Artificial Intelligence Research Symposium*, pp. 295–299.

Zadeh LA, 1986. "Is probability theory sufficient for dealing with uncertainty in AI: A negative view" in Kanal LN and Lemmer JF (eds), *Uncertainty in Artificial Intelligence*. North-Holland, pp. 103–116.

Zhong N, Dong J-Z and Ohsuga S, 1997. "Discovering rules in the environment with noise and incompleteness" *Proceedings of the Tenth International Florida Artificial Intelligence Research Symposium*, pp. 186–191.