Artificial Intelligence 59 (1993) 39–42
Elsevier

ARTINT 985

Probabilistic logic revisited

Nils J. Nilsson

Robotics Laboratory, Department of Computer Science, Stanford University, Stanford, CA 94305, USA

1. Origins

Before beginning the research that led to "Probabilistic logic" [11], I had participated with Richard Duda, Peter Hart, and Georgia Sutherland on the PROSPECTOR project [3]. There, we used Bayes' rule (with some assumptions about conditional independence) to deduce the probabilities of hypotheses about ore deposits given (sometimes uncertain) geologic evidence collected in the field [4]. At that time, I was also familiar with the use of "certainty factors" by Shortliffe [18], the use of "fuzzy logic" by Zadeh [20], and the Dempster/Shafer formalism [16]. All of these methods made (sometimes implicit and unacknowledged) assumptions about underlying joint probability distributions, and I wanted to know how the mathematics would work out if no such assumptions were made. I began by asking how *modus ponens* would generalize when one assigned probabilities (instead of binary truth values) to P and $P \supset Q$. As can be verified by simple calculations using a Venn diagram, the probability of Q is under-determined in this case but can be bounded as follows:

$$p(P) + p(P \supset Q) - 1 \leqslant p(Q) \leqslant p(P \supset Q).$$

The techniques that I developed in the paper for calculating bounds on probabilities can be understood as a kind of generalization of the Venn diagram method. While I was working out the ideas in "Probabilistic logic", I was unaware of similar work by Good [7], Smith [19], and de Finetti [2].

Given probabilities on sentences, one can do no better than calculate bounds on derived sentences because the probabilities of the given sentences

Correspondence to: N.J. Nilsson, Robotics Laboratory, Department of Computer Science, Stanford University, Stanford, CA 94305, USA. E-mail: nilsson@cs.stanford.edu.

do not, in general, completely determine the underlying joint distribution. However, it is of interest to ask about the minimum-entropy joint distribution because, perhaps unlike some of the other methods for reasoning with uncertain information, the minimum-entropy distribution assumes minimal additional information. For this reason, I included in the paper Cheeseman's technique for minimum entropy [1].

My primary aim in writing "Probabilistic logic" was to present an intuitively reasonable but foundational account of the problem of uncertain reasoning. The complete impracticability of calculating the bounds prescribed in the paper was of little concern to me because I imagined that approximate methods might later be devised, and I even suggested an approximate method in the paper. Pearl [12, p. 463] mentions that a mechanism in Quinlan's INFERNO [14] can be regarded as a local approximation to probabilistic logic.

2. Main contribution

The key intellectual contribution of "Probabilistic logic" was a formal procedure for calculating the bounds on the probability of a sentence in the predicate calculus given the probabilities (or the bounds on the probabilities) of other sentences. I called this process "probabilistic entailment" because it is based on models of the sentences. The contribution served mainly to elucidate the foundations of probabilistic reasoning even though it is in general, intractable. I hoped that it would set the stage for possible approximate methods and for comparison with other methods.

3. Open issues

A major omission from the paper was any discussion of proof-theoretic methods for making probabilistic deductions. The paper stimulated some attempts to develop deductive techniques; see, for example, recent work by Haddawy and Frisch who have found a complete set of inference rules for a subset of probabilistic logic [9].

Devising good approximate methods for probabilistic entailment is still an important area for future work. It would seem that assumptions about conditional independence (as might be represented by influence diagrams or belief networks) could be used to simplify the bounds calculations in probabilistic entailment. In that connection, see a recent paper by Fertig and Breese [6].

Judea Pearl [13] has persuaded me that I should have shown more explicitly how probabilistic logic should handle assignments of conditional rather than absolute probabilities. For example, a more natural generalization of *modus ponens* emerges if we specify $p(P)$ and, then, $p(Q|P)$ instead of $p(P \supset Q)$. The probability of Q is then bounded by

$$p(Q|P)p(P) \leqslant p(Q) \leqslant 1.$$

Pearl [12, p. 459] argues that the probability $p(P \supset Q)$ does not properly reflect what we normally mean by the certainty of the rule "if P then Q". For example, if we want to say that some rare event P has a likely consequence Q, and we write $p(P) = 0.01$ and $p(P \supset Q) = 0.9$, we find that the two sentences are inconsistent. Writing $p(P) = 0.01$ and $p(Q|P) = 0.9$ gives the bound $0.09 \leqslant p(Q) \leqslant 1$, which is more reasonable.

4. Subsequent work

Work on probabilistic reasoning exploded after the mid-1980s. Several "Workshops on Uncertainty in Artificial Intelligence" have been held, and their proceedings have been published. The volume edited by Shafer and Pearl contains many important papers as well as illuminating perspectives by the editors about various aspects of uncertain reasoning [17]. Pearl has also written an indispensable text on the subject [12], now available in a revised and updated second printing.

Fagin and Halpern have presented a more formal and general analysis than that contained in my paper [5].

Two other important developments have occurred. One is the use of belief networks and influence diagrams to represent causal relations that allow simplifying assumptions to be made about the conditional independence of propositions (see [12, Chapters 2–4] and [15]). Another development involves analyses by Heckerman [10] and by Grosof [8] that establish important connections among techniques that use (slightly modified) certainty factors, the Dempster/Shafer formalism, and special cases of Bayes' rule.

5. Conclusions

Since my foray into probabilistic reasoning was brief, and because I have not acquainted myself with the extensive and growing literature in this field, I will refrain from attempting any sage remarks about how my work might relate to that of others. The reader might want to consult a few paragraphs by Pearl, however, on the circumstances under which one might want to use probabilistic logic [12, pp. 461–462]. I am also indebted to Judea Pearl for his comments and suggestions about the present note.

References

[1] P.Cheeseman, A method of computing generalized Bayesian probability values for expert systems, in: *Proceedings IJCAI-83*, Karlsruhe, Germany (1983)

[2] B. de Finetti, *Theory of Probability* (Wiley, New York, 1974).

[3] R.O. Duda, The PROSPECTOR system for mineral exploration, Final Report to the U.S. Geological Survey, Contract No. 14-08-0001-17296, SRI International, Menlo Park, CA (1980).

[4] R.O. Duda, P.E. Hart and N.J. Nilsson, Subjective Bayesian methods for rule-base inference systems, in: *Proceedings 1976 National Computer Conference*, AFIPS **45** (1976) 1075–1082; also in: B.W. Webber and N.J. Nilsson, eds., *Readings in Artificial Intelligence* (Morgan Kaufmann, San Mateo, CA, 1981).

[5] R. Fagin and J. Halpern, Uncertainty, belief, and probability, in: *Proceedings IJCAI-89*, Detroit, MI (1989) 1161–1167.

[6] K.W. Fertig and J.S. Breese, Interval influence diagrams, in: M. Henrion, R.D. Shachter, T.S. Levitt, L.N. Kanal and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence* **5** (Elsevier Science Publishers B.V., Amsterdam, 1990) 149–161.

[7] I.J. Good, *Probability and the Weighing of Evidence* (Griffin, London, 1950).

[8] B. Grosof, Evidential confirmation as transformed probability, in: L.N. Kanal and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence* (North-Holland, Amsterdam, 1986).

[9] P. Haddawy and A. Frisch, Anytime deduction for probabilistic logic, Unpublished Manuscript (submitted for publication).

[10] D. Heckerman, Probabilistic interpretations for MYCIN's certainty factors, in: L.N. Kanal and J.F. Lemmer, eds., *Uncertainty in Artificial Intelligence* (North-Holland, Amsterdam, 1986) 167–196; also in: G.A. Shafer and J. Pearl, *Readings in Uncertain Reasoning* (Morgan Kaufmann, San Mateo, CA, 1990) 298–312.

[11] N.J. Nilsson, Probabilistic logic, *Artif. Intell.* **28** (1986) 71–87.

[12] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, San Mateo, CA, 1988; revised and updated second printing, 1991).

[13] J. Pearl, Private communication (1992).

[14] J.R. Quinlan, INFERNO: a cautious approach to uncertain inference, *Comput. J.* **26** (1983) 255–269.

[15] R.D. Shachter, Evaluating influence diagrams, *Oper. Res.* **33** (6) (1986); also in: G.A. Shafer and J. Pearl, *Readings in Uncertain Reasoning* (Morgan Kaufmann, San Mateo, CA, 1990) 79–90.

[16] G.A. Shafer, *Mathematical Theory of Evidence* (Princeton University Press, Princeton, NJ, 1979).

[17] G.A. Shafer and J. Pearl, *Readings in Uncertain Reasoning* (Morgan Kaufmann, San Mateo, CA, 1990).

[18] E.H. Shortliffe, *Computer-Based Medical Consultations: MYCIN* (American Elsevier, New York, 1976).

[19] C.A.B. Smith, Consistency in statistical inference and decision, *J. R. Stat. Soc. Ser. B* **23** (1961) 218–258.

[20] L.A. Zadeh, Fuzzy logic and approximate reasoning, *Synthese* **30** (1975) 407–428.