DISTRIBUTED INFERENCING FOR CLASSIFICATION

FINAL PROJECT REPORT

N00014-89-J-2007, 1/1/89 - 6/30/91

Judea Pearl
Department of Computer Science
University of California, Los Angeles, CA 90024

1. OVERVIEW

This objective of this project was to develop algorithms, tools, and architectures that will improve the classification of complex systems. The approach was based on representing the system as a network of parallel, autonomous units, communicating via messages that summarize the state of neighboring subsystems. This approach offers reduction in complexity, ease of programming, universality, and potential for parallel hardware implementation. A major part of the effort focused on augmenting the traditional belief-network representation with non-Bayesian methods, qualitative information, variable-strength defaults statements, and causal interactions.

Our main effort has concentrated on seeking non-Bayesian methods for classification, where we have explored the applicability of Ordinal Conditional Functions [1] and the Dempster-Shafer belief functions [2]. The need for such formalisms stems from the fact that we often do not possess the probabilistic knowledge required for full Bayesian analysis. For example, we may not know the components' failure rates and, which is more often the case, we may not be able to rate which among several possible modes of failure is likely to be realized when a given component fails.

We have concluded that such ignorance cannot be adequately represented in the Dempster-Shafer formalism and, moreover, the applications of belief functions methods can lead to implausible test and replacement strategies. Our studies show that the belief measure *Bel* computes the probability that a given component can be *proven* faulty, rather than the probability that it *is* faulty [3]. This shift in semantics leads to unpredictable and sometimes implausible behavior and it appears that strategies based on these measures cannot be implemented safely [2].

Our evaluation of Ordinal Conditional Functions has been more encouraging, and it led to a major effort toward a QUALITATIVE formulation of probabilistic diagnosis. The reasons for taking this approach are several. First, when an expert does not have a precise knowledge of the probabilities involved in the diagnostic system a coarse qualitative estimate of these probabilities would often be adequate. Second, the qualitative formulation can utilize the machinery of symbolic processing, thus offering a computational advantage coupled with a more natural explanation facility.

The approach we have taken draws upon the interpretation of linguistic quantifiers (e.g., likely, very likely, etc,) as encodings of infinitesimal probabilities, to be processed and combined according to the axioms of probability theory, and to be translated back to linguistic form. The result is a probabilistically sound calculus, employing integer addition, for manipulating order-of-magnitudes of disbeliefs. For example, if we make the following correspondence between linguistic expressions and ε^n :

$P(A) = \varepsilon^0$	A is believable	$\kappa(A) = 0$
$P(A) = \varepsilon^1$	A is unlikely	$\kappa(A) = 1$
$P(A) = \varepsilon^2$	A is very unlikely	$\kappa(A) = 2$
$P(A) = \varepsilon^3$	A is extremely unlikely	$\kappa(A) = 3$
----	------------------------	----

then the infenitesimal approximation yields a nonmonotonic logic to reason about likelihood. For example, it can take sentences in the form of quantified conditional sentences, e.g., "Birds are likely to fly", (written $\kappa(\neg f \mid b) = 1$), "Penguins are most likely birds", (written $\kappa(\neg b \mid p) = 2$), "Penguins are extremely unlikely to fly," (written $\kappa(f \mid p) = 3$) and return quantified conclusions in the form of "If x is a penguin-bird then x is extremely unlikely to fly" (written $\kappa(f \mid p \wedge b) = 3$).

The basic κ ranking system, as described in Spohn [1] requires the specification of a complete probability model before reasoning can commence. In other words, the knowledge base must be sufficiently rich to define the κ associated with every world w. In practice, such specification might require knowledge that is not readily available in common discourse. For example, we might be given the information that birds fly (written $\kappa(\neg f \mid b) = 1$) and no information at all about properties of non-birds, thus leaving $\kappa(f \wedge \neg b)$ unspecified. Hence, stronger inferential machinery is required for drawing conclusions from partially specified models, like those associating a κ with isolated default statements.

Such machinery is provided by a calculus called System-Z^+, which we have developed in the past year [3]. The calculus admits fragmentary sets of conditional sentences, treats them as constraints over the distribution of $\kappa(w)$, and infers only such statements that are compelled to acquire high likelihood in the "most normal" distribution $\kappa(w)$ satisfying these constraints.

In the second phase of the project, we explored the applicability of the qualitative reasoning provided by System-Z^+ to problems of diagnosis, with special emphasis on distributed diagnosis. In principle, System-Z^+ should enjoy a similar diagnostic power as that of traditional Bayesian networks, the aim being to find a set S of faults that minimizes $\kappa^+(S \mid \text{evidence})$. This is indeed the case when we possess sufficient knowledge to specify a complete κ^+ distribution and to embed its dependence structure in a network form. However, when knowledge is only partial, special techniques have been developed for mixing structural knowledge (e.g. independence constraints) with mechanism knowledge (i.e., conditional likelihood constraints). One such method, based on causal stratification of rules, promises to provide the qualitative counterpart of Bayesian networks.

References

[1] Spohn, W., A General non-probalistic theory of inductive reasoning. In: *Proceedings of the 4th Workshop on Uncertainty in AI,* Minneapolis, MN, 1988, 315-322.

[2] Pearl, J., "Reasoning with belief functions: An analysis of compatibility," *International Journal of Approximate Reasoning,* Vol. 4, (5/6), 1990, 363-389.

[3] Pearl, J., *Probabilistic reasoning in intelligent systems,* San Mateo, CA: Morgan Kaufmann, 1988.

[4] Goldszmidt, M. and Pearl, J., System-Z$^+$: A formalism for reasoning with variable-strength defaults. In: *Proceedings, AAAI-91*, Anaheim, CA, July 1991, 399-404.

2. SPECIFIC RESULTS

The following results were obtained during the period of performance:

-- A new representation of graphoids in the form of "multi-graphs" was developed and analyzed (R-133, R-126, R-127).

--The application of graphical models in statistical analysis were characterized and categorized using properties of graphoids (R-97, R-114, R-142 , R-116).

-- Polynomial-time algorithms were developed to test the consistency of assertions about dependencies and independencies (R-119).

-- The application of graphoids to structure learning have resulted in a universal learning methods for certain graph structures and causal models (R-152).

-- New formulation of "relevance" relations were developed, with applications to the acquisition of probabilistic knowledge (R-141, R-142).

-- A new theory of default reasoning was developed, combining the benefits of both the extensional and conditional interpretations (R-157).

-- A tractable procedure was developed for testing the consistency of databases containing a mixture of defeasible and strict sentences (R-122, R-131, R-139).

-- Effective methods were developed for reasoning with inheritance networks, embodying specificity, defeat, chaining and causality (R-129, R-137).

-- Algorithms were developed for default reasoning based on the maximum-entropy principle (R-144).

-- Methods of identifying useful structures in empirical data were formulated and analyzed (R-132, R-142, R-149 R-152).

-- A comparative study was conducted comparing Bayesian and Belief-Functions strategies in automated reasoning (R-136, R-143).

-- An empirical definition of causation was formulated, and sound algorithms for discovering causal structures in statistical data were developed (R-150, R-155, R-156).

-- The role of hidden variables in constraint network was analyzed and distributed algorithms for solving the network consistency problem were developed (R-147, R-148).

-- A relational semantics was developed for directed (causal) constraint networks, and algorithms for organizing constraints in causal structures were devised (R-153).

-- New formalisms for reasoning with variable-strength defaults were developed and shown to be tractable (R-122, R-154, R-161).

-- A propositional semantics and tractable algorithms were developed for subsets of default logic. (R-163, R-169, R-170).

-- A method of interpreting and processing nonmonotonic causal rules was developed (R-171).

3. PUBLICATIONS

List of publications resulting from the NRL Award.

1. **(R-97)**. Geiger, D. & Pearl, J., "Logical and Algorithmic Properties of Conditional Independence." A short version (R-97-II-S) in *Proceedings, 2nd International Workshop on Artificial Intelligence and Statistics*, Miami, Florida, January 1989, pp. (19-1)-(19-10). A long version (R-97-II-L), forthcoming, *The Annals of Statistics*.

2. **(R-114)**. Pearl, J., Geiger, D. & Verma, T., "The Logic of Influence Diagrams." *Proceedings, Conference on Influence Diagrams for Decision Analysis, Inference, and Prediction*, Berkeley, California, May 1988. Also in R.M. Oliver and J.Q. Smith (Eds), *Influence Diagrams, Belief Nets and Decision Analysis*, Sussex, England: John Wiley & Sons, Ltd., 1989, pp. 67-87. A shorter version (R-114-S), in *Kybernetika*, Vol. 25:2, 1989, pp. 33-44.

3. **(R-116)**. D. Geiger, Verma, T.S. & Pearl, J., "Identifying Independence in Bayesian Networks." *Networks*, 1990, Vol. 20, No. 5, pp. 507-534.

4. **(R-119)**. D. Geiger, A. Paz & J. Pearl, "Axioms and Algorithms for Inferences Involving Probabilistic Independence." *Information and Computation*, Vol. 91, No. 1, March 1991, pp. 128-141.

5. **(R-122)**. M. Goldszmidt & J. Pearl, "Deciding Consistency of Databases Containing Defeasible and Strict Information." *Proceedings, 5th Workshop on Uncertainty in AI*, Windsor, Ontario, Canada, August 1989, pp. 134-141. Also in *Uncertainty in Artificial Intelligence*, 5, M. Henrion, R.D. Shachter, L.N. Kanal, and J.F. Lemmer (Eds.). Elsevier Science Publishers B.V. (North Holland), 1990, pp. 87-97. Expanded version, "On the consistency of defeasible databases," *Artificial Intelligence*, North-Holland Publishers, Amsterdam, (52), 121-149, December 1991.

6. **(R-123)**. D. Geiger & J. Pearl, "Logical and Algorithmic Properties of Independence and Their Application to Bayesian Networks." *Annals of Mathematics and AI*, Vol. 2, No. 1-4, pp. 165-178, 1990.

7. **(R-126)**. Geva, R.Y., & Paz. A., "Towards Complete Representation of Graphoids in Graphs." *Proceedings of the 15th International Workshop on Graph Theoretic Concepts in Computer Sciences*, Rodluc, 1989. Lecture Notes in Computer Sciences, No. 411, pp. 41-62, Springer Verlag.

8. **(R-127)**. Paz, A., "The Theory of Graphoids - A Survey." *Proceedings of the International Workshop on Graph-Grammars*, Bremen, 1990. To appear as a Springer Lecture Notes in Computer Sciences.

9. **(R-128)**. J. Pearl, "Probabilistic Semantics for Nonmonotonic Reasoning: A Survey." *Proceedings, First Intl. Conf. on Principles of Knowledge Representation and Reasoning*, Toronto, Canada, May 1989, pp. 505-516, Expanded version in *Philosophical AI: Computational Approaches to Reasoning*, ed. Robert Cummins and John Pollock, Bradford Books/MIT Press, 157-187, 1991.

10. **(R-129)**. H. Geffner & T. Verma, "Inheritance = Chaining + Defeat." A condensed version (R-129-S) in *Methodologies for Intelligent Systems, IV.*, edited by Z. Ras. New York: Elsevier Science Publishing Co., Fall 1989, pp. 411-418.

11. **(R-130)**. D. Geiger, T. Verma & J. Pearl, "*d*-Separation: From Theorems to Algorithms." *Proceedings, 5th Workshop on Uncertainty in AI*, Windsor, Ontario, Canada, August 1989, pp. 118-124. Also in *Uncertainty in AI*, 5, M. Henrion, R.D. Shachter, L.N. Kanal, & J.F. Lemmer (Eds.) Elsevier Science Publishers (North Holland), 1990, pp. 139-148.

12. **(R-131)**. J. Pearl, "System Z: A Natural Ordering of Defaults with Tractable Applications to Nonmonotonic Reasoning." *Theoretical Aspects on Reasoning about Knowledge*, R. Parikh (ed), San Mateo: Morgan Kaufmann, 1990, pp. 121-135.

13. **(R-132)**. J. Pearl & R. Dechter, "Learning Structure from Data: A Survey." *Proceedings, 2nd Workshop on Computational Learning Theory (COLT'89)*, Santa Cruz, CA, August 1989, pp. 230-244.

14. **(R-133)**. Ur. S., & Paz, A., "The Representation Power of Probabilistic Knowledge by Undirected Graphs and Directed Acyclic Graphs: A Comparison." *Proceedings of the Workshop on Uncertainty Processing in Expert Systems*, Alsovice Czechoslovakia, 1991.

15. **(R-134)**. Paz, A. "Graphoids." *Encyclopedia of AI, Second Ed.*, John Wiley & Sons, Inc., New York, 592-593, 1992.

16. **(R-136).** J. Pearl, "Reasoning with Belief Functions: An Analysis of Compatibility." *The International Journal of Approximate Reasoning*, Vol. 4, No. 5/6, 1990, pp. 363-389.

17. **(R-137).** H. Geffner, "Default Reasoning: Causal and Conditional Theories." UCLA Cognitive Systems Laboratory, *Technical Report (R-137), CSD #890065*, November 1989. PhD. Dissertation, MIT Press, 1991.

18. **(R-139).** Goldszmidt, M. & Pearl, J., "On the Relation Between Rational Closure and System-Z." *Proceedings, Third International Workshop on Nonmonotonic Reasoning*, So Lake Tahoe, CA, May 31-June 3, 1990, pp. 130-140.

19. **(R-141).** D. Geiger & D. Heckerman, "Practical and Theoretical Advances in Knowledge Acquisition of Probabilistic Networks." *Proceedings, Sixth Conference on Uncertainty in AI*, Cambridge, Massachusetts, July 27-29, 1990 (entitled "Separable and Transitive Graphoids"), pp. 538-545.

20. **(R-142).** D. Geiger, "Graphoids: A Qualitative Framework for Probabilistic Inference." UCLA Cognitive Systems Laboratory, *Technical Report (R-142)*, January 1990. PhD. Dissertation.

21. **(R-143).** J. Pearl, "Which is More Believable, The Probably Provable or the Provably Probable?" *Proceedings, CSCSI-90, Eighth Canadian Conference on Artificial Intelligence*, Ottawa, CA, May 23-25, 1990, pp. 1-7.

22. **(R-144).** M. Goldszmidt, P. Morris, & J. Pearl, "A Maximum Entropy Approach to Nonmonotonic Reasoning." *Proceedings, AAAI-90*, Boston, MA, 7/29-8/3/90, pp. 646-652.

23. **(R-145).** J. Pearl, "Probabilistic and Qualitative Abduction." *Proceedings, AAAI Spring Symposium on Abduction*, Stanford, March 27-29, 1990, pp. 155-158.

24. **(R-147).** R. Dechter, "On the Expressiveness of Networks with Hidden Variables." *Proceedings, AAAI-90*, Boston, MA, 7/29-8/3/90, pp. 556-562.

25. **(R-148).** Z. Collin & R. Dechter, "A Distributed Solution to the Network Consistency Problem." *Proceedings, ISMIS-90*, Lennox, TN.

26. **(R-149).** D. Geiger, A. Paz & J. Pearl. "Learning Causal Trees from Dependence Information." *Proceedings, AAAI-90*, Boston, MA, 7/29-8/3/90, pp. 770-776.

27. **(R-150).** T. Verma & J. Pearl. "Equivalence and Synthesis of Causal Models." *Proceedings, Sixth Conference on Uncertainty in AI*, Cambridge, Mass., July 27-29, 1990, pp. 220-227.

28. **(R-152).** D. Geiger, A. Paz and J. Pearl. "Learning Simple Causal Networks." Forthcoming, *International Journal of Intelligent Systems*, John Wiley & Sons.

Inc., Vol. 8, 1993.

29. **(R-153)**. R. Dechter and J. Pearl. "Directed Constraint Networks: A Relational Framework for Causal Modeling." *Proceedings, 12th International Joint Conference of Artificial Intelligence (IJCAI-91)*, Sydney, Australia. August 24-30, 1991, pp. 1164-1170.

30. **(R-154)**. M. Goldszmidt and J. Pearl. "Where do Default Priorities Come From?" UCLA Cognitive Systems Laboratory, Computer Science Dept. *Technical Report (R-154)*, November, 1990.

31. **(R-155)**. J. Pearl and T. Verma. "A Statistical Semantics for Causation." *Proceeding, 3rd International Workshop on A.I. & Statistics*, Fort Lauderdale, Florida, January 2-5, 1991.

32. **(R-156)**. J. Pearl and T. Verma. "A Theory of Inferred Causation." In Allen, J.A., Fikes, R., and Sandewall, E. (Eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the Second International Conference*. San Mateo, CA: Morgan Kaufmann. (April, 1991), pp. 441-452.

33. **(R-157)**. H. Geffner and J. Pearl. "Conditional Entailment: Bridging Two Approaches to Default Reasoning." *Artificial Intelligence*, 53, 209-244, 1992.

34. **(R-159)**. J. Pearl. "Epsilon-Semantics." *Encyclopedia of AI, Second Edition*, John Wiley & Sons, Inc., New York, 468-475, 1992.

35. **(R-160)**. I. Meiri. "Combining Qualitative and Quantitative Constraints in Temporal Reasoning." *Proceedings of AAAI-91*, Anaheim, CA, July 1991, Vol. 1, pp. 260-267.

36. **(R-161)**. M. Goldszmidt and J. Pearl. "System-Z^{+}: A formalism for reasoning with variable-strength defaults." *Proceedings, AAAI-91*, Anaheim, CA, July 1991, Vol. 1, pp. 399-404.

37. **(R-162)**. R. Ben-Eliyahu and R. Dechter. "Default Logic, Propositional Logic and Constraints." *Proceedings, AAAI-91*, Anaheim, CA, July 1991, Vol. 1, pp. 379-385.

38. **(R-163)**. R. Ben-Eliyahu and R. Dechter. "Propositional Semantics for Default Logic." *Proceedings of the Fourth International Workshop on* Nonmonotonic Reasoning, Plymouth, VT, 13-27, May 1992.

39. **(R-164)**. N. C. Dalkey. "Entropy and Belief Networks." *Proceedings of the 8th Conference on Uncertainty in Artificial* Intelligence, Stanford, CA, Morgan Kaufmann, San Mateo, CA, 55-58, July 1992.

40. **(R-167)**. Z. Collin, R. Dechter, and S. Katz. "On the Feasibility of Distributed

Constraint Satisfaction'' *Proceedings, IJCAI'91*, Sydney, Australia, 318-324, August 24-30, 1991.

41. **(R-168)**. J. Pearl, R. Dechter, and T. S. Verma. ''Knowledge Discovery VS. Data Compression.'' *Proceedings, KDD-91 Workshop*, 207-211, 1991.

42. **(R-170-S)**. R. Ben-Eliyahu and R. Dechter. ''Propositional Semantics for Disjunctive Logic Programs.'' UCLA, Cognitive Systems Laboratory, *Technical Report (R-170)*, October 1991.

43. **(R-171)**. M. Goldszmidt and J. Pearl. ''Stratified Rankings for Causal Modeling.'' *Proceedings of the Fourth International Workshop on* Nonmonotonic Reasoning, Plymouth, VT, 99-110, May 1992.

4. STUDENTS

List of students supported by the NRL Award.

Alex Balke

Rachel Ben-Eliyahu

Huy Cao

Hector Geffner (PhD, 1989) ''Default Reasoning: Causal and Conditional Theories.'' Co-winner ACM Best Dissertation Award

Dan Geiger (PhD, 1990) ''Graphoids: A Qualitative Framework for Probabilistic Inference.''

Moises Goldszmidt (PhD, 1992), ''Qualitative Probabilities: A Normative Framework for Commonsense Reasoning.''

Jim Kan

Itay Meiri (PhD, 1992), ''Temporal Reasoning: A Constraint-Based Approach.''

Sek-Wah Tan

Thomas Verma (degree expected: PhD, Winter 1993)

Amir Weinshtain (MS, 1991)