UNCERTAINTY IN INTELLIGENT SYSTEMS

edited by

Bernadette Bouchon-Meunier
LAFORIA
University of Paris VI
Paris, France

Llorenc Valverde
Department of Information Sciences
University of the Balears
Palma(Balears), Spain

Ronald R. Yager
Machine Intelligence Institute
Iona College
New Rochelle, NY, U.S.A.

1993

NORTH-HOLLAND
AMSTERDAM • LONDON • NEW YORK • TOKYO

Uncertainty in Intelligent Systems
B. Bouchon-Meunier et al. (Editors)

Valuation Networks, Decision Trees, and Influence Diagrams: A Comparison

Prakash P. Shenoy

School of Business, University of Kansas, Summerfield Hall, Lawrence, KS 66045-2003, USA
pshenoy@ukanvm.cc.ukans.edu

1. INTRODUCTION

Recently, we proposed a new method for representing and solving Bayesian decision problems based on the framework of valuation-based systems [Shenoy 1992b, 1993a]. The new representation is called a valuation network, and the new solution method is called the fusion algorithm. In this paper, we briefly compare valuation networks to decision trees and influence diagrams. For symmetric decision problems, valuation networks are more expressive than both decision trees and influence diagrams. We also compare the fusion algorithm to the backward recursion method of decision trees and to the arc-reversal method of influence diagrams. For symmetric decision problems, the fusion algorithm is more efficient than the backward recursion method of decision trees, and more efficient and simpler than the arc-reversal method of influence diagrams.

An outline of this paper is as follows. In section 2, we give a statement of the *Medical Diagnosis* problem. In section 3, we describe a decision tree representation and solution of the *Medical Diagnosis* problem highlighting the strengths and weaknesses of the decision tree method. In section 4, we describe an influence diagram representation and solution of the *Medical Diagnosis* problem highlighting the strengths and weaknesses of the influence diagram technique. While the strengths of the influence diagram technique are well known, its weaknesses are not so well known. For example, it is not well known that in data-rich domains where we have a non-causal graphical probabilistic model of the uncertainties, influence diagrams representation technique is not very convenient to use. Also, the inefficiency of the arc-reversal method has never before been shown in the context of decision problems. In section 5, we describe a valuation network representation and solution of the *Medical Diagnosis* problem highlighting its strengths and weaknesses vis-á-vis decision trees and influence diagrams. For more details of the comparison, see [Shenoy 1993b].

2. A MEDICAL DIAGNOSIS PROBLEM

A physician is trying to determine a policy for treating patients suspected of suffering from a disease D. D causes a pathological state P that in turn causes symptom S to be exhibited. The physician first observes whether or not a patient is exhibiting symptom S. Based on this observation, she either treats the patient (for D and P) or does not. The physician's utility function depends on her decision to treat or not, the presence or absence of disease D, and the presence or absence of pathological state P. The prior probability of disease D is 10%. For patients known to suffer from D, 80% suffer from pathological state P. On the other hand, for patients known not to suffer from D, 15% suffer from P. For patients known to suffer from P, 70% exhibit symptom S. And for patients known not to suffer from P, 20% exhibit symptom S. We assume D and S are conditionally independent given P. Table 1 shows the physician's utility function.

Table 1. The physician's utility function for all act-state pairs.

Physician's Utilities (υ)		States			
		Has pathological state (p)		No pathological state (~p)	
		Has disease (d)	No disease (~d)	Has disease (d)	No disease (~d)
Acts	Treat (t)	10	6	8	4
	Not treat (~t)	0	2	1	10

3. DECISION TREE REPRESENTATION AND SOLUTION

A popular method for representing and solving Bayesian decision problems is decision trees. Decision trees have their genesis in the pioneering work of von Neumann and Morgenstern [1944] on extensive form games. Decision trees graphically depict all possible scenarios. The decision tree representation allows computation of an optimal strategy by the backward recursion method of dynamic programming [Zermelo 1913, Bellman 1957, Raiffa and Schlaifer 1961]. Raiffa and Schlaifer [1961] call the dynamic programming method for solving decision trees "averaging out and folding back." Figure 1 shows a decision tree representation of the *Medical Diagnosis* problem. Figure 2 shows its solution.

The strengths of decision trees are its simplicity and its flexibility. Decision trees are based on the semantics of scenarios. Each path in a decision tree from the root to a leaf represents a scenario. These semantics are very intuitive and easy to understand. Decision trees are also very flexible. In asymmetric decision problems, the choices at any time and the relevant uncertainty at any time depend on past decisions and revealed events. Since decision trees depict scenarios explicitly, representing an asymmetric decision problem is easy.

The weaknesses of decision trees are its modeling of uncertainty, its modeling of information constraints, and its combinatorial explosiveness in problems in which there are many variables. Since decision trees are based on the semantics of scenarios, the placement of a random variable in the tree depends on the information constraints. Also, decision trees demand a probability distribution for each random variable conditioned on the past decisions and events leading to the random variable in the tree. This is a problem in diagnostic decision problems where we have a causal model of the uncertainties. For example, in medical decision problems, symptoms are revealed before diseases. For such problems, decision trees require conditional probabilities for diseases given symptoms. But, assuming a causal model, it is easier to assess the conditional probabilities of symptoms given the diseases [Shachter and Heckerman 1987]. Thus a traditional approach is to first assess the probabilities in the causal direction and then compute the probabilities required in the decision tree using Bayes theorem. This is a major drawback of decision trees. There should be a cleaner way of separating a representation of a problem from its solution. The former is hard to automate while the latter is easy. Decision trees interleave these two tasks making automation difficult. This drawback of decision trees can be alleviated by using information sets [see Shenoy 1993d].

Figure 1. A decision tree representation of the *Medical Diagnosis* problem.

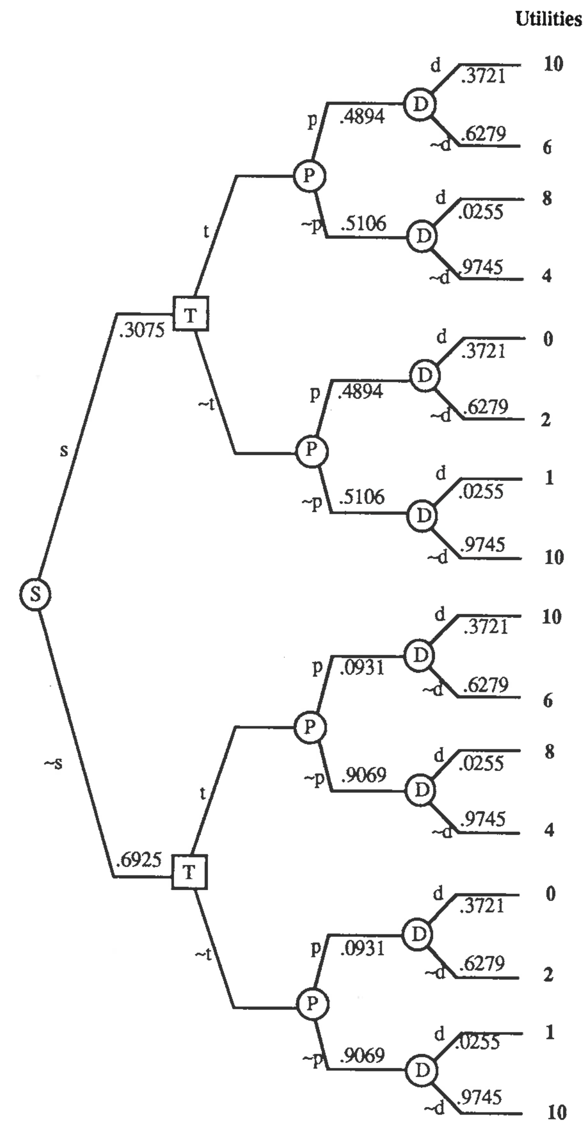
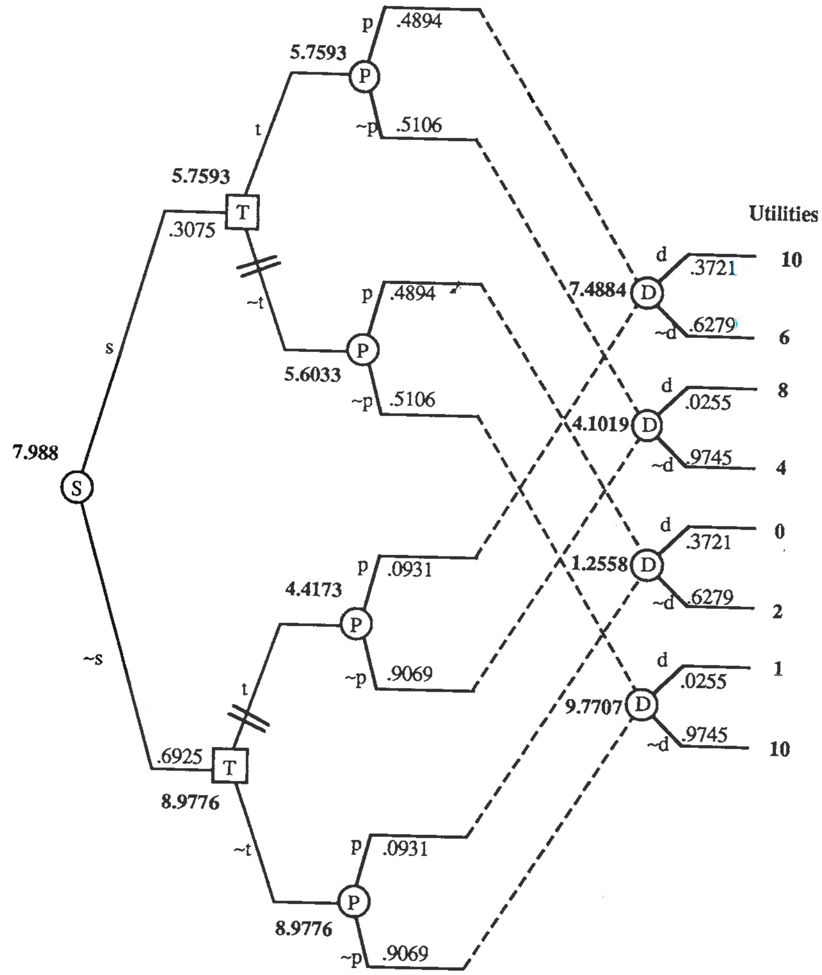


Figure 2. A decision tree solution of the *Medical Diagnosis* problem using coalescence.



In decision trees, the sequence in which the variables occur in each scenario represents information constraints. In some problems, the information constraints may only be specified up to a partial order. But the decision tree representation demands a complete order. This overspecification of information constraints in decision trees makes no difference in the final solution. However, it may make a difference in the computational effort required to compute a solution.

The combinatorial explosiveness of decision trees stems from the fact that the number of scenarios is an exponential function of the number of variables in the problem. In a symmetric decision problem with n variables, where each variable has 2 possible values, there are 2^n

scenarios. Since decision trees depict all scenarios explicitly, it is computationally infeasible to represent a decision problem with, say, 50 variables.

The strength of the decision tree solution procedure is its simplicity. Also, if a decision tree has several identical subtrees, then we can make the solution process more efficient by coalescing the subtrees [Matheson and Roths 1967].

The weakness of the decision tree solution procedure is the preprocessing of probabilities that may be required (before the decision tree representation). A brute-force computation of the desired conditionals from the joint distribution for all variables is intractable if there are many random variables. Also, although preprocessing is required for representing the problem as a decision tree, some of the resulting computations are unnecessary for solving the problem [Shenoy 1993b].

4. INFLUENCE DIAGRAM REPRESENTATION AND SOLUTION

Influence diagram is another method for representing and solving decision problems. Influence diagrams were initially proposed as a method only for representing Bayesian decision problems [Miller et al. 1976, Howard and Matheson 1981]. A motivation behind the formulation of influence diagrams was to find a method for representing decision problems without any preprocessing. Subsequently, Olmsted [1983] and Shachter [1986] devised methods for solving influence diagrams directly, without first having to convert influence diagrams to decision trees. In the last decade, influence diagrams have become popular for representing and solving decision problems [Oliver and Smith 1990]. Figure 3 shows an influence diagram representation of the *Medical Diagnosis* problem. Figure 4 shows its solution.

Figure 3. An influence diagram for the *Medical Diagnosis* problem.

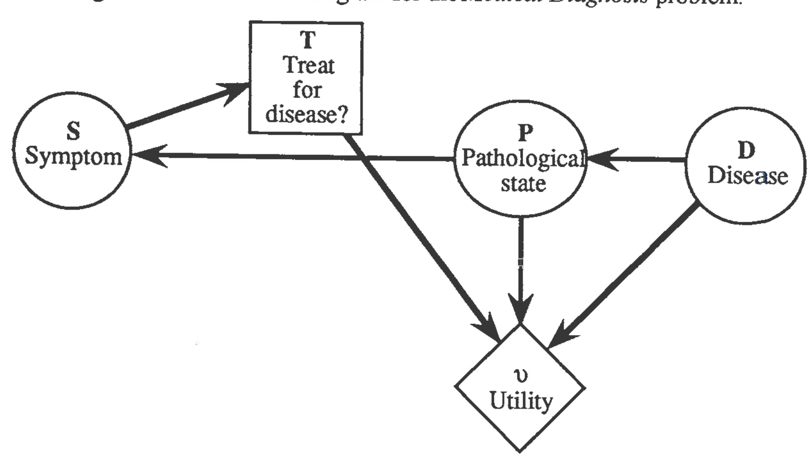
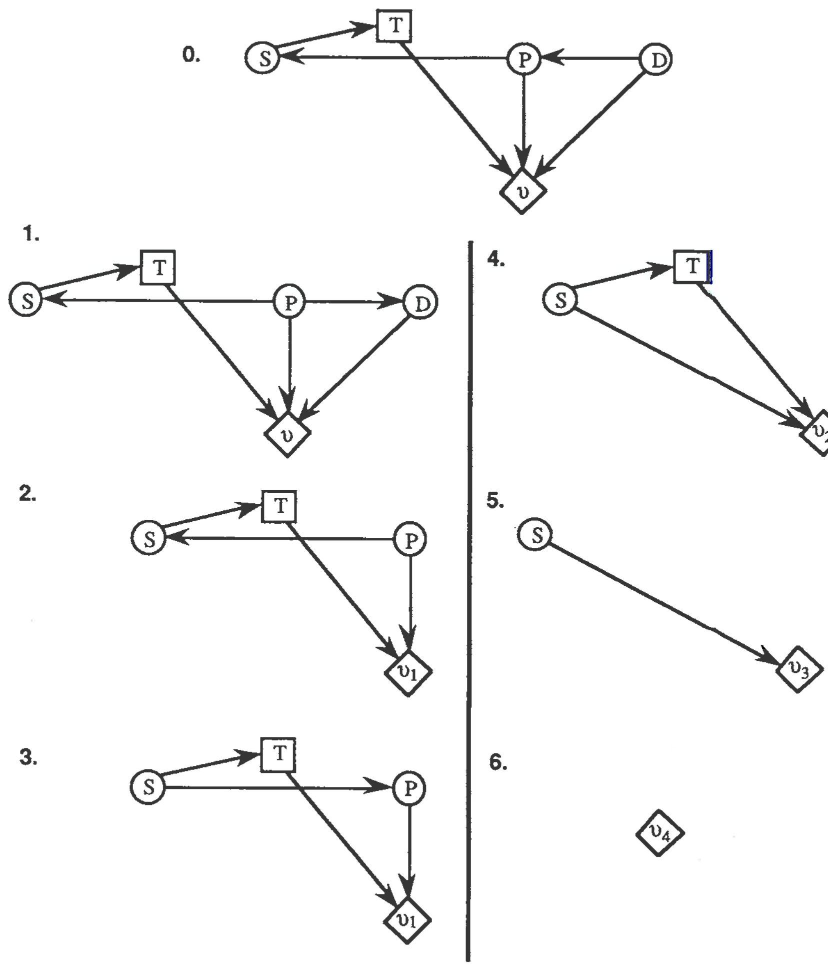


Figure 4. The solution of the influence diagram representation of the *Medical Diagnosis* problem.



The strengths of the influence diagram representation are its modeling of uncertainty and its compactness. Influence diagrams are based on the semantics of conditional independence. Conditional independence is represented in influence diagrams by d-separation of variables [Pearl et al. 1990]. Practitioners who have used influence diagrams in their practice claim that it is a powerful tool for communication, elicitation, and detailed representation of human knowledge [Owen 1984, Howard 1988, 1989, 1990].

Influence diagrams do not depict scenarios explicitly. They assume symmetry (i.e., every scenario consists of the same sequence of variables) and depict only the variables and the sequence up to a partial order. Therefore, influence diagrams are compact and computationally more tractable than decision trees.

The weaknesses of the influence diagram representation are its modeling of uncertainty and requirement of symmetry. Influence diagrams demand a conditional probability distribution for each random variable. In causal models, these conditionals are readily available. However, in other graphical models, we don't always have the joint distribution expressed in this way [see, e.g., Darroch et al. 1980, Wermuth and Lauritzen 1983, Edwards and Kreiner 1983, and Kiiveri et al. 1984]. For such models, before we can represent the problem as an influence diagram, we have to preprocess the probabilities, and often, this preprocessing is unnecessary for the solution of the problem [Shenoy 1993b].

Influence diagrams are suitable only for decision problems that are symmetric or almost symmetric [Watson and Buede 1987]. For decision problems that are highly asymmetric, influence diagram representation is awkward. For such problems, Call and Miller [1990], Fung and Shachter [1990], Covaliu and Oliver [1992], and Kirkwood [1993] investigate representations that are hybrids of decision trees and influence diagrams, and Olmsted [1983], and Smith et al. [1989] suggest methods for making the influence diagram representation more efficient.

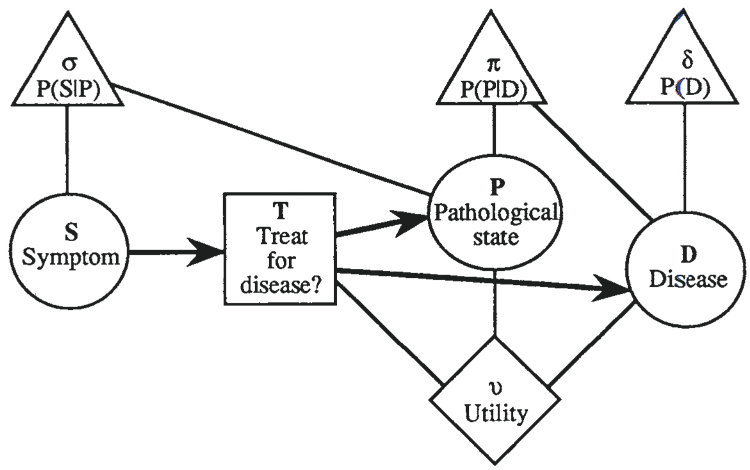
The strength of the influence diagram solution procedure is that, unlike decision trees, it uses local computation to compute the desired conditionals in problems requiring Bayesian revision of probabilities [Shachter 1988, Rege and Agogino 1988, Agogino and Ramamurthi 1990]. This makes possible the solution of large problems in which the joint probability function decomposes into small functions.

The weakness of the arc-reversal method for solving influence diagrams is that it does unnecessary divisions. The solution process of influence diagrams has the property that after deletion of each variable, the resulting diagram is an influence diagram. As we have already mentioned, the representation method of influence diagrams demands a conditional probability distribution for each random variable in the diagram. It is this demand for conditional probability distributions that requires divisions, not any inherent requirement in the solution of a decision problem.

5. VALUATION NETWORK REPRESENTATION AND SOLUTION

Valuation network is yet another method for representing and solving Bayesian decision problems [Shenoy 1992b, 1993a]. Like influence diagrams, valuation networks depict decision variables, random variables, utility functions, and information constraints. Unlike influence diagrams, valuation networks explicitly depict probability functions. Valuation networks are based on the semantics of factorization. Each probability function is a factor of the joint probability distribution function, and each utility function is a factor of the joint utility function. The solution method for valuation networks is called the fusion algorithm. Figure 5 shows a valuation network representation of the *Medical Diagnosis* problem. Figure 6 shows its solution.

Figure 5. A valuation network for the *Medical Diagnosis* problem.



The strengths of valuation networks are its expressiveness for modeling uncertainty, its compactness, and its simplicity. Unlike decision trees and influence diagrams, valuation networks do not demand specification of the joint probability distribution function in a certain form. All probability models can be represented directly without any preprocessing.

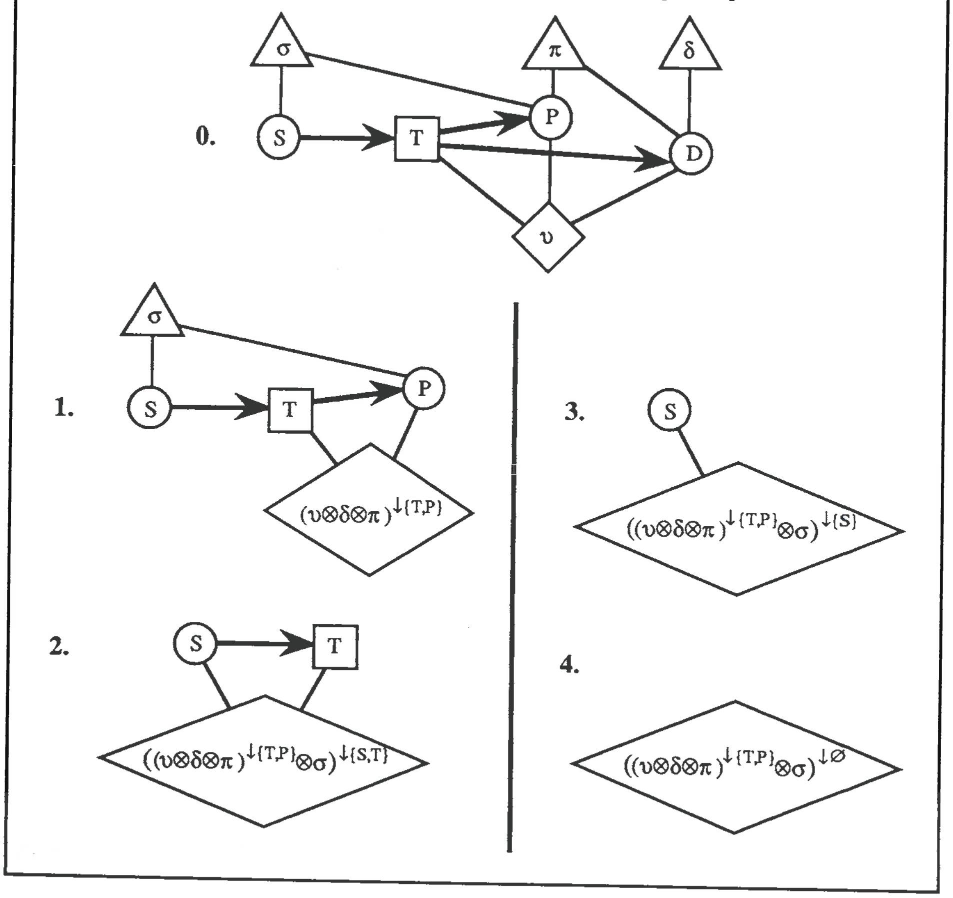
Like influence diagrams, valuation networks are compact representations of decision problems. They do not depict scenarios explicitly. Only variables, functions, and information constraints are depicted explicitly. Of course, like influence diagrams, valuation networks assume symmetry of scenarios.

Valuation networks are very simple to interpret. Each probability function is a factor of the joint probability distribution function, and each utility function is a factor of the joint utility function. Thus it treats probability functions and utility functions alike. Given the factors of the joint probability distribution, we can easily recover the independence conditions underlying the joint probability distribution using separation of variables [Shenoy 1993c]. Also, the information constraints are represented explicitly.

A weakness of the valuation network representation is that, like influence diagrams, it is appropriate only for symmetric and almost symmetric decision problems. For highly asymmetric decision problems, the use of valuation networks is awkward. Another weakness of valuation networks is that the semantics of factorization are not as well developed as the semantics of conditional independence. Hopefully, current research on the semantics of factorization will address this shortcoming of valuation networks [Shenoy 1989, 1991, 1992a, 1992c].

The strengths of the fusion algorithm for solving valuation networks are its computational efficiency and its simplicity. The fusion algorithm uses local computations, and it avoids unnecessary divisions. This makes it always more efficient than the arc-reversal method of influence diagrams. It is also more efficient than the backward recursion method of decision trees for symmetric decision problems. The fusion algorithm is also extremely simple to understand and execute.

Figure 6. The fusion algorithm for the *Medical Diagnosis* problem.



The weakness of the fusion algorithm is that, in the worst case, its complexity is exponential. This is not surprising because solving a decision problem is NP-hard [Cooper 1990]. In problems with many variables, the fusion algorithm is tractable only if the sizes of the frames on which combinations are done stay small. The sizes of the frames on which combinations are done depend on the sizes of the domains of the functions and on the information constraints. We need strong independence conditions to keep the sizes of the potentials small. And we need strong assumptions on the joint utility function to decompose it into small functions. Of course, this weakness is also shared by decision trees and influence diagrams.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation under grant SES-9213558. I am grateful for discussions with and comments from Ali Jenzarli, Pierre Ndilikilikesha, Anthony Neugebauer, Geoff Schemmel, and Leen-Kiat Soh.

REFERENCES

Agogino, A.M. and K. Ramamurthi. 1990. Real time influence diagrams for monitoring and controlling mechanical systems. In Oliver, R. M. and J. Q. Smith (eds). *Influence Diagrams, Belief Nets and Decision Analysis*. 199–228. John Wiley & Sons, Chichester.

Bellman, R. E. 1957. *Dynamic Programming*. Princeton University Press, Princeton, NJ.

Call, H. J. and W. A. Miller. 1990. A comparison of approaches and implementations for automating decision analysis. *Reliability Engineering and System Safety*. **30**, 115–162.

Cooper, G. F. 1990. The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence*. **42**, 393–405.

Covaliu, Z. and R. M. Oliver. 1992. Formulation and solution of decision problems using decision diagrams. Unpublished manuscript. University of California at Berkeley, CA.

Darroch, J. N., S. L. Lauritzen, and T. P. Speed. 1980. Markov fields and log-linear interaction models for contingency tables. *The Annals of Statistics*. **8**, 522–539.

Edwards, D. and S. Kreiner. 1983. The analysis of contingency tables by graphical models. *Biometrika*. **70**, 553–565.

Ezawa, K. J. 1986. Efficient evaluation of influence diagrams. Ph.D. dissertation. Department of Engineering-Economic Systems, Stanford University.

Fung, R. M. and R. D. Shachter. 1990. Contingent influence diagrams. Unpublished manuscript. Advanced Decision Systems, Mountain View, CA.

Howard, R. A. 1988. Decision analysis: Practice and promise. *Management Science*. **34**(6), 679–695.

Howard, R. A. 1989. Knowledge maps. *Management Science*. **35**(8), 903–922.

Howard, R. A. 1990. From influence to relevance to knowledge. In Oliver, R. M. and J. Q. Smith (eds.). *Influence Diagrams, Belief Nets and Decision Analysis*. 3–23. John Wiley & Sons, Chichester.

Howard, R. A. and J. E. Matheson. 1981. Influence diagrams. Reprinted in Howard, R. A. and J. E. Matheson (eds.). 1984. *The Principles and Applications of Decision Analysis*. **2**, 719–762. Strategic Decisions Group, Menlo Park, CA.

Kiiveri, H., T. P. Speed, and J. B. Carlin. 1984. Recursive causal models. *Journal of the Australian Mathematics Society, Series A*. **36**, 30–52.

Kirkwood, C. W. 1993. An algebraic approach to formulating and solving large models for sequential decisions under uncertainty. *Management Science*. To appear.

Matheson, J. E. and W. J. Roths. 1967. Decision analysis of space projects: Voyager Mars. Reprinted in R. A. Howard and J. E. Matheson (eds.). 1984. *Readings on The Principles and Applications of Decision Analysis*. **1**, 446–475. Strategic Decisions Group, Menlo Park, CA.

Mellouli, K. 1987. On the propagation of beliefs in networks using the Dempster-Shafer theory of evidence. Ph.D. dissertation. School of Business, University of Kansas, Lawrence, KS.

Miller III, A. C., M. W. Merkhofer, R. A. Howard, J. E. Matheson, and T. R. Rice. 1976. Development of Decision Aids for Decision Analysis. Final Technical Report DO #27742. Stanford Research Institute, Menlo Park, CA.

Ndilikilikesha, P. 1991. Potential influence diagrams. Working Paper No. 235. School of Business, University of Kansas.

Oliver, R. M. and J. Q. Smith (eds.). 1990. *Influence Diagrams, Belief Networks, and Decision Analysis*. John Wiley & Sons, Chichester.

Olmsted, S. M. 1983. On representing and solving decision problems. Ph.D. dissertation. Department of Engineering-Economic Systems, Stanford University.

Owen, D. L. 1978. The use of influence diagrams in structuring complex decision problems. Reprinted in Howard, R. A. and J. E. Matheson (eds.). 1984. *Readings on The Principles and Applications of Decision Analysis*. **2**, 765–772. Strategic Decisions Group, Menlo Park, CA.

Pearl, J., D. Geiger, and T. Verma. 1990. The logic of influence diagrams. In Oliver, R. M. and J. Q. Smith (eds.). *Influence Diagrams, Belief Nets and Decision Analysis*. 67–88. John Wiley & Sons, Chichester.

Raiffa, H. 1968. *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*. Addison-Wesley, Reading, MA.

Raiffa, H. and R. Schlaifer. 1961. *Applied Statistical Decision Theory*. MIT Press, Cambridge, MA.

Rege, A. and A. M. Agogino. 1988. Topological framework for representing and solving probabilistic inference problems in expert systems. *IEEE Transactions on Systems, Man, and Cybernetics*. **18**(3), 402–414.

Shachter, R. D. 1986. Evaluating influence diagrams. *Operations Research*. **34**, 871–882.

Shachter, R. D. 1988. Probabilistic influence diagrams. *Operations Research*. **36**, 589–604.

Shachter, R. D. and D. E. Heckerman. 1987. A backwards view for assessment. *AI Magazine*. **8**(3), 55-61.

Shenoy, P. P. 1989. A valuation-based language for expert systems. *International Journal of Approximate Reasoning*. **3**(5), 383–411.

Shenoy, P. P. 1991. Conditional independence in valuation-based systems. Working Paper No. 236. School of Business, University of Kansas, Lawrence, KS.

Shenoy, P. P. 1992a. Valuation-based systems: A framework for managing uncertainty in expert systems. In Zadeh, L. A. and J. Kacprzyk (eds.). *Fuzzy Logic for the Management of Uncertainty*. 83–104. John Wiley & Sons, New York, NY.

Shenoy, P. P. 1992b. Valuation-based systems for Bayesian decision analysis. *Operations Research*. **40**(3), 463–484.

Shenoy, P. P. 1992c. Conditional independence in uncertainty theories. In D. Dubois, M. P. Wellman, B. D'Ambrosio and P. Smets (eds.). *Uncertainty in Artificial Intelligence: Proceedings of the Eighth Conference*. 284–291. Morgan Kaufmann, San Mateo, CA.

Shenoy, P. P. 1993a. A new method for representing and solving Bayesian decision problems. In Hand, D. J. (ed.). *Artificial Intelligence Frontiers in Statistics: AI and Statistics III*. 119–138. Chapman & Hall, London.

Shenoy, P. P. 1993b. A comparison of graphical techniques for decision analysis. *European Journal of Operational Research*. To appear.

Shenoy, P. P. 1993c. Valuation networks and conditional independence. Working Paper No. 238. School of Business, University of Kansas, Lawrence, KS.

Shenoy, P. P. 1993d. Game trees for decision analysis. Working Paper No. 239. School of Business, University of Kansas, Lawrence, KS.

Smith, J. E., S. Holtzman, and J. E. Matheson. 1989. Structuring conditional relationships in influence diagrams. Unpublished manuscript. Fuqua School of Business, Duke University.

Tatman, J. A. and R. D. Shachter. 1990. Dynamic programming and influence diagrams. *IEEE Transactions on Systems, Man, and Cybernetics*. **20**(2), 365–379.

von Neumann, J. and O. Morgenstern. 1944. *Theory of Games and Economic Behavior*. 1st edition. John Wiley & Sons, New York, NY.

Watson, S. R. and D. M. Buede. 1987. *Decision Synthesis: The Principles and Practice of Decision Analysis*. Cambridge University Press, Cambridge, U.K.

Wermuth, N. and S. L. Lauritzen. 1983. Graphical and recursive models for contingency tables. *Biometrika*. **70**, 537–552.

Zermelo, E. 1913. Uber eine anwendung der mengenlehre auf die theorie des schachspiels. *Proceedings of the Fifth International Congress of Mathematics*. **2**, 501–504. Cambridge, U.K.