Value of Information Analysis in Dynamic Decision Models

Songsong Xu Kim-Leng Poh
Department of Industrial & Systems Engineering
National University of Singapore
Kent Ridge, Singapore 119260
{ engp8920, isepohkl }@nus.edu.sg

Abstract

Decision-making under dynamic environments is an important problem frequently encountered in the practice. In this paper, we reveal some properties of value of information in dynamic environment, and compare the results with that in different representations by other researchers. We study the time invariant features in dynamic influence diagrams (DID) to enhance the value of information calculation in dynamic decision models. We identify the structure in DIDs which could be decomposed to temporal invariant sub-DIDs and then be constructed into sub-junction trees to improve the efficiency in calculation of the total expected value and the value of information. We also discuss issues including finding spreading variables which will complicate the calculation, adding mapping variables directly into the sub-junction trees to enable a simple calculation for value of information for decision-intervening variables, and discounting benefits among the sub-junction trees according to time. Finally, we propose an algorithm based on these discussions which is polynomial to the size of the network when there are no spreading variables in the DID.

1 Introduction

An action taken at a certain time point may affect the decision maker's current revenue, and his/her future benefits through the influence of this action on other actions to be taken some time in the future. Knowing the value of information for some uncertain variables in the decision problems is quite desirable to direct the information gathering procedure to improve the quality of decision in such dynamic environments.

Value of information has been well defined and deeply studied in general influence diagrams. The idea of economic evaluation of information in decision making was first introduced by Howard (1966, 1967). Raiffa (1968) classical textbook described an exact method for computing the expected value of perfect information (EVPI).

In recent years, there has been great interest in developing schemes for computing the value of information. Exact methods for computing the value of information have been explored (Ezawa, 1994; Howard and Matheson, 1981; and Shacher, 1990). Unfortunately, just like the general inference, the computational complexity of such exact computation of EVPI in a general decision model with any general utility function is known to be intractable. Even with the simplifying assumptions that a decision maker is risk neutral or has a constant degree of risk aversion, the problem remains intractable.

The intractability of EVPI computation has motivated researchers to explore a variety of quantitative approximations, including myopic, iterative one-step lookahead procedures (Gorry, 1973; Heckerman, Horvitz & Nathwani, 1992; Dittmer and Jensen, 1997; Shachter, 1999) and nonmyopic procedures based on employing arguments hinging on the statistical law for large numbers (Heckerman, Horvitz & Middleton, 1991). Furthermore, several works on qualitative evaluation of EVPI (Poh & Horvitz, 1996) and hybrid method for combining both qualitative and quantitative approaches (Xu, Poh & Horvitz, 2001) have also been proposed to accelerate the EVPI calculation for important uncertain variables.

In many practical cases the decision problems of similar structures appear over and over again, hence we address such problems as dynamic decision models. However, few studies have been focused on the VOI calculation in dynamic environment to take advantage of the time-invariant features though the general influence diagram has a sequence of decisions in nature.

A lot of representations can be adopted to model the dynamic decision problems. Tatman and Shachter (1990) extended influence diagrams into dynamic influence diagrams (DIDs) by allowing time-separable value functions, one for each time unit or decision stage. We adopt this representation for the inheritance in concepts and algorithms. They can also be viewed as other dynamic optimization problems, e.g., partially observable Markov decision processes (POMDPs).

We present in this paper some properties of VOI in dynamic environment, and compare the results with that of other researchers in different representations. For time-invariant DIDs, we also identify a group of DIDs which can be decomposed into sub-networks with similar structures according to time, and hence a sub-junction tree can be generated as computing template. We address the opinions on reusing the original junction tree to calculate VOI, even for decision-intervening variables. We also address the time value in the computation. Finally we propose an algorithm based on the previous discussions that is polynomial if the DID has no spreading variables.

2 VOI in DIDs

First let's consider a naïve case of decision maker takes a set of actions A blindly, without knowing anything about the system status S, as shown in Figure 1, where D, X, V represent the decision, the chance and the utility variable respectively.

The case of observing the system state all the time is the same as completely observable Markov decision processes, like figure 2, which have the classical dynamic programming (value iteration) equation (Howard, 1960).

The EVPI thus can be obtained as the difference between expected utility in the two figures.

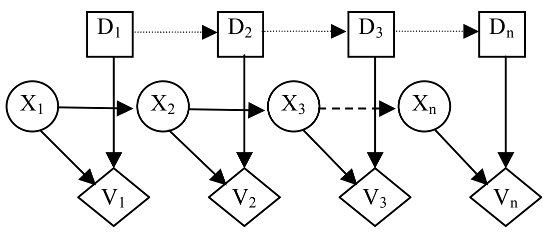

Figure 1: A naïve DID without information

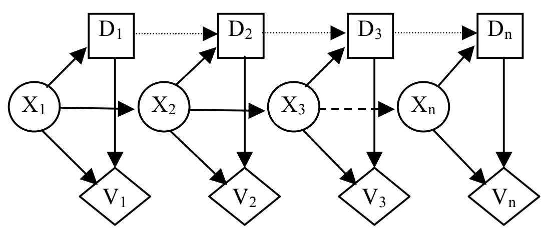

Figure 2: DID with perfect information (MDP)

Denote the total value of knowing nothing, knowing all and having only delayed information as V^0, V' and V^1 respectively. We have proved theoretically that $V^0 \leq V^1 \leq V'$, (Xu and Poh, 2002). This means it is always preferable to observe the system state earlier than later, and later than never if costs are not concerned. The difference of value between observation before an early decision Dj and a later one D_k can be called the VOI of temporal delay.

This inequality of expected value for different information availability is meaningful, because if the VOI ordering were in contrary case, there would be no need to concern the opportunity cost occurred in information gathering period. Since observing earlier is better than later, the problem becomes a tradeoff when facing the opportunity cost.

3 Temporally Invariant Junction Tree for DIDs

Kjærulff (1992) proposed the Dynamic Expansion and Reduction (DER) method to perform exact inference in Dynamic Bayesian Networks (DBNs) by adding new time slices and deleting old ones dynamically. However, in many practical cases the system structure in every time slice is stationary or near stationary. Xiang (1999) argued that pre-compiling some slice representation of a stationary DBN could support more efficient exact inference. A sub-junction tree template was constructed from the original DBN by first identifying a subnet, $S_{i-1} \cup N_i \cup S_i$, where S_{i-1} and S_i are minimal separators of the DBN and N_i is the part between them, as shown in Figure 3.

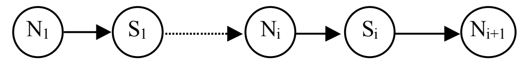

Figure 3: Partition of a DBN

In order to calculate value of information in dynamic systems, we need to deal with the decisions and values. Shachter (1999) presented an algorithm based on Bayes-Ball (Shachter, 1998) to determine the requisite observations of each decision and then change the influence diagram into a belief network in time linear to the graph size. Though it was developed in a decision system with separable value nodes to prune the set of information predecessors for each of the decisions in the influence diagram, it didn't take full advantage of the repeatability of most practical dynamic systems.

3.1 Problem

Our objective is trying to apply efficient clustering method in a DID, making use of the stationary or near stationary features of the system. We hope to build a template junction tree and then evolve the system dynamically to calculate the value of information. The procedure is: first identify a subnet to build template junction tree; then apply Bayes-Ball algorithm to convert the decision problem into a probabilistic network; construct template junction tree thereafter; evolved into the next time stage by updating the current belief and finally reuse the junction tree to calculate the value of information.

However, we may not be able to include all the requisite observations in the template subnet with their corresponding decisions. For example, in Figure 4, $V_n \perp (D_n, I_n)$, $I_n = \{b_0, D_0, ..., b_{n-1}, D_{n-1}, b_n\}$, the requisite observation set of D_n, $R_n = \{b_n, D_{n-1}\}$. Applying the Requisite Observation Algorithm we find that $R_i = \{b_0, ..., b_i, D_{n-1}\}$. We must add many arcs from these requisite observations to the decision, as shown in Figure 5. Thus the resulting belief network is much more complex, and no longer Markovian.

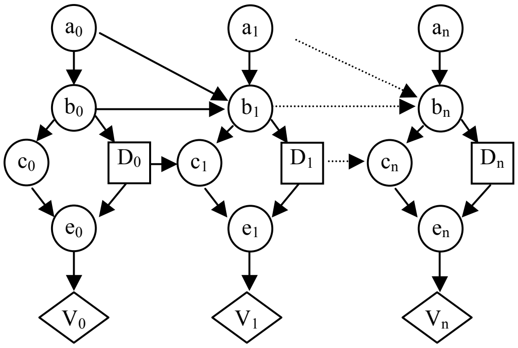

Figure 4: An example of DID

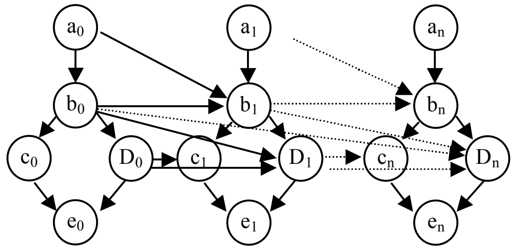

Figure 5: Resulting DBN for the example above

This is because the information predecessor b_i has a parent a_i which is also included in the forward interface. The ball passed from b_{i+1} bounce back to b_i from a_i, and bounce back to b_i's other parents that are in the previous stage. According to the repeating structural character of the temporally invariant DIDs, the Bayes-Ball will stop at the previous stage thus the requisite observations of the decision are all included in a small subnet; otherwise it will pass to the very first stage and cause the structure like what is shown in Figure 5.

Hence before we divide the whole DID into subnets, a check in the structure to ensure the DID is decomposable is needed. If the DID cannot be decomposed to time-invariant subnets, it is not likely that we can save much computational time and space by constructing template and rolling back thereafter.

If a node can carry the Bayes-Ball to the previous stages without being blocked, thus spread the requisite observations to those stages, we call it a *spreading variable*.

Spreading variables:

In a DID $M = \{X, D, E, V\}$, a node $X_i \notin I_i$, we can show that if X_i is **not** *d*-separated to W_{i+1} by the set $W_i \cup D_i$, then X_i is a spreading variable.

Induction.

Start from the last stage n, run the Bayes-Ball algorithm on $(V_n \mid D_n, I_n)$, the requisite observation R_n should lie in $W_n \cup D_n$.

For the n-1th stage, run the Bayes-Ball algorithm on $(V_{n-1} \cup R_n \mid D_{n-1}, I_{n-1})$. If W_n is not *d*-separated from X_n by $W_{n-1} \cup D_{n-1}$, and W_n is not *d*-separated from X_{n-1} by $W_{n-1} \cup D_{n-1}$, the ball from R_n can pass to X_{n-1} through some active path, and pass to W_{n-1} if the network is connected, which is a trivial constraint. Reason by analogy, the ball will be passed from W_i to W_{i-1} and then to the very first stage through the active paths. Thus the requisite observation set will include variables in all the previous stages.

If there are any spreading variables in the DID, it is hard to take the advantage of the time invariant features of the dynamic model. Otherwise, all the requisite observations are located within the subnet N_i, which is from $W_{i-1} \cup D_{i-1}$ to $W_i \cup D_i$. To be exact, the subnet is between the forward interfaces of the DBN after converted from DID. We can thus identify the requisite observation of each decision in such subnet for the value node and the requisite observations obtained in the later subnet, without running the Bayes-Ball Algorithm over the entire DID. After requisite observations are identified, this subnet might differ from the original sub-ID on structure with arcs from requisite observations added to the decision and arcs from non-requisite observations removed, but it inherits the temporal invariant feature. The forward interface (Xiang, 1999) based on this subnet will also be a self-sufficient separator and the sub-junction tree constructed be properly constructed with a root cluster that has no children.

The condition for the absence of spreading variables, $(X_{i+1} \perp X_i \mid W_i, D_i)$, means the variables in the current stage are independent of the history given the decisions and their information variables in the previous stage. This is not a very strict restriction, e.g., the common Markov Decision Processes have such feature, and hence they can be solved using dynamic programming iteratively. Our DID is more than simple MDP in that it includes many unobservable variables, as long as it satisfies the condition.

However, some Partially Observable Markov Decision Processes which the state variables are dependent on the whole history, thus cannot be separated by the decisions and the observations before those decisions, are not temporally decomposable. In these cases, the largest clique constructed will increase as the number of the time stages t increases, namely at least t in clique size. We show this procedure as the following: if a decision D_i has the requisite observations refer back to previous stages, e.g., nodes $B_0, \ldots, B_{i-1}$, then add arcs from $B_0, \ldots, B_{i-1}$ to D_i, and also from $B_0, \ldots, B_{i-2}$ to D_{i-1}, etc. In moralization, every pair of B_k and B_j ($k \neq j \in \{0, \ldots, i\text{-}1\}$) are linked with a moral arc, hence the nodes $B_0, \ldots, B_{i-1}$ and D_i form a complete set of size i. Reason by analogy, if there are n sets of such requisite information for D_i, $A_0, \ldots, A_{i-1}, B_0, \ldots, B_{i-1}, \ldots$, then the clique-width will increase to at least nT. The clique includes these requisite observations and their child decision node is the largest cluster in the junction tree corresponds to the DID.

The other parts without spreading variables in them can still be triangulated and constructed as clusters with the same structure separately, in a recursive manner.

The other problem we address here is the re-using of the junction tree to calculate the value of information on the basis of the junction tree template. Dittmer et al (1997) reused the original junction tree to calculate the value of observing a variable *A* before *D* by adding *A* to all the clusters between *A* and *D*'s inward-most cliques. This method is feasible in the dynamic environment since the conclusion is over an ID with multiple decisions and separable values. If a particular variable is to be observed several stages earlier, the sub-junction tree with this variable added in the cliques can be reused.

3.2 Mapping Variables

Sometimes we would like to know the VOI for the decision-intervening chance nodes. However, an arc cannot be directly added from the decision to the chance node under discussion since this forms a loop in the original influence diagram, which means a deadlock in calculation. This is addressed by adding mapping variables and thus forcing influence diagrams to be formulated (or reformulated) into canonical form (Howard 1990).

We find that adding mapping variables in a specific stage only influences the construction of the sub- junction tree for the subnet of that stage. When the distance between the decision and the chance node is not far, usually it only adds some leaf clusters including the mapping variables to the junction tree.

The procedure of adding mapping variables to convert an ID to canonical form has been introduced in Heckerman & Shachter (1995).

Denote the responsive chance node as X_i, and the unresponsive chance nodes which are its parents as $\Pi(X_i)$. The mapping variables are added between X_i and $\Pi(X_i)$. Before the mapping variable $X_i(C_i)$ is added, X_i is in the same clique of $\Pi(X_i)$. With the mapping variable added this clique splits to two, one includes $X_i(C_i)$ and $\Pi(X_i)$, and the other includes X_i and $X_i(C_i)$, as shown in the right part of Figure 6. If X_i has no chance node parents, only a leaf clique including X_i, $X_i(C_i)$, and C_i is added.

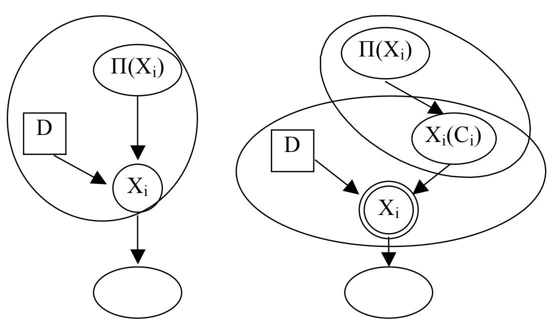

Figure 6: ID without or with mapping variable added

In a properly constructed junction tree, a decision *D* is weakly inward of uncertainty *A* if *A* is a descendant of *D* in the influence diagram, hence it is easy to check if *D* is weakly inward of *A*, and determine if the mapping variable of *A* are necessary to add or not based on the junction tree, if a junction tree is already built for the sake of inference.

3.3 Discounting

In dynamic systems time value is considered, so the benefits and received and the costs incurred in each decision stage should be discounted with time.

Discounting takes place when a discount factor ρ is applied to the value node merging operation in DIDs. Given the net values (benefit with cost deducted) $r_0, \ldots, r_n$ for each stage, the merged value of *i*th stage $V_i = r_i + \rho V_{i+1}$.

The discounting has not yet been discussed in the junction tree constructed from an influence diagram. In Shachter (1999), if the value nodes are not nested, which is the case of dynamic influence diagram (time-separable value nodes), the structure of properly constructed junction tree will be as the Figure 7, for each un-nested value node, Q_i is the non-value variables relevant to V_i. Denote the utility potential as ψ, and probability potential as ϕ. To each clique *C* in the junction tree, for chance variable *X*, the marginalization operation is $\mathop{\mathbf{M}}_X = \sum_X \phi_C$; and for decision variable *D*, $\mathop{\mathbf{M}}_D = \max_D \phi$. If *T* is a junction tree, C_1 and C_2 are adjacent cliques with separator $S \subset T$, and if $C_1 \prec C_2$, then C_1 absorbs from C_2, the value / utility potential ψ_{Ci} is propagated from leaves to root, if it is not null. Hence a discount factor $\rho \in (0, 1)$ can be directly multiplied to the second addend of the equation to denote the time value when the marginalization proceeds in reversed time order:

$$\phi'_{C_1} = \phi_{C_1} * \phi_S; \qquad \psi'_{C_1} = \psi_{C_1} + \rho \cdot \frac{\psi_S}{\phi_S}$$

here,

$$\phi_S = \mathop{\mathbf{M}}_{C_2 \setminus S} \phi_{C_2}; \qquad \psi_S = \mathop{\mathbf{M}}_{C_2 \setminus S} \phi_{C_2} * \psi_{C_2}$$

Typically the economic interest rate is used for such discounting of both cost and benefit. However, the choice of discounting rate has been disputed by researchers; a rate of 2-10% is often considered as consistent with economic theory (Petitti, 1994).

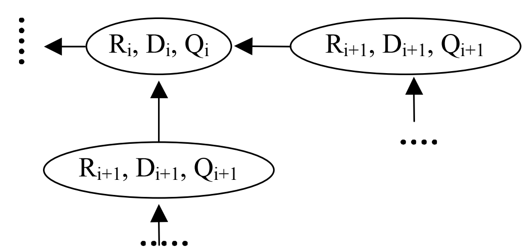

Figure 7: A part of properly constructed junction tree

4 Calculating VOI in Dynamic Influence Diagrams

Input: A temporally invariant dynamic influence diagram *M* (2-stage representation and number of stages *T*), the chance node X_i of interest, the decision node D_i prior to which we observe X_i.
Output: The value of information of observing X_i before D_i.
Procedure:

1. Check if there are any spreading variables in the DID.
2. If no, specify the subnet between the forward interfaces.
   a. Find requisite observations R_i for $(V_i \mid D_i, I_i)$ using the Bayes-Ball algorithm. Run Bayes-Ball on $(V_{i-1} \cup R_i \mid D_{i-1}, I_{i-1})$
   b. Store all the requisite observations in R_i, make R_i as only parents of D_i.
   c. For $i = 0$ to T do
      Construct sub-junction tree
      Calculate the expected value of sub-ID
      Discount and update the potentials
3. If yes, change the whole DID into a DBN. Construct junction tree, and calculate the discounted expected value.
4. Add mapping variables if necessary. Add X_i to all cliques inward of the clique with D_i, re-calculate the expected value and finally get the difference.

Above is the proposed algorithm for calculating VOI in DIDs according to the earlier discussions. The Bayes-Ball algorithm runs in time linear to the graph size. To judge if there are any spreading variables, there's no need to run it in the whole DID, but two stages of the DID. Hence if a time-decomposable DID has *T* time stages with *N* nodes in each stage, the time will be $O(N^2)$. The time for loading the probabilities and performing the calculations is proportional to the total space given by $\Sigma_{C\in H} \Pi_{v\in C} |S(v)|$, where *C* is the clique in the junction tree *H* and *v* is a vertex in *C*, *S* is the state space of *v*. This is dominated by the size of the maximal clique if all the vertices have the same state space size. The inference time is $O(T \cdot N \cdot \prod_{Xi\in C} S(X_i))$ where *C* is the largest clique in the junction tree.

Becker and Geiger (1996) proposed an approximate algorithm to find a near optimal (errs by a factor *a*, which is 3.66 in the paper) junction tree in polynomial time. Thus our inference will be $O\{T \cdot N \cdot poly(N) \cdot [\prod_{Xi\in C} S(X_i)]^a\}$, where *poly*(*N*) is the complexity for linear programming, usually polynomial. The time for running Bayes-Ball is negligible, compare to the inference time, hence the total time will be polynomial.

For the case that spreading variables exist, the inference is much more complex. If the clique-width is slightly greater than the logarithm of *N*, there is no polynomial algorithm unless P equals to NP. As we discussed earlier, the clique-width grows with the decision stages. Solving such problems exactly will be hard, no matter what is the representation.

5 Conclusions

We address in this paper the problem of value of information calculation in dynamic decision models, namely dynamic influence diagrams. We reveal some properties of VOI in dynamic environment, and compare the results with that in different representations by other researchers.

We also identify a group of DIDs which can be decomposed into sub-networks with similar structures, and hence a sub-junction tree can be generated based on such sub-networks as the computing template. We discuss the methods on reusing the sub-junction tree when VOI for variables with intervening decisions is under concern. We consider the time value of benefits and costs by discounting the value in each stage.

We propose an algorithm of calculating VOI in dynamic decision models. When there are no spreading variables in the DID, the algorithm is polynomial to the network. However, when spreading variables exist, the system is very complex and calculating VOI is hard for most exact inference approaches so far.

The method is supposed to provide researchers a tool of sensitive analysis in dynamic decision modeling. We believe it is both illustrative and efficient.

Future research might focus on the efficient way to find an optimal or near optimal triangulation, and calculating VOI for a group of variables simultaneously and apply the method to real world case.

References:

Beinlich, I. A., Suermondt, H. J., Chavez, R. M. & Cooper, G. F. 1989. The ALARM Monitoring System: A Case Study with Two Probabilistic Inference Techniques for Belief Networks. *Second European Conference on Artificial Intelligence in Medicine*, London, 38:247-256.

Becker, A, Geiger, D. 1996. A sufficiently fast algorithm for finding close to optimal junction trees, *Proceedings of the Twelfth Uncertainty in Artificial Intelligence*, Morgan Kaufmann.

Dittmer, S., L. and Jensen, F., V. 1997. Myopic Value of Information in Influence Diagrams. In D. Geiger and P. P. Shenoy (eds.), *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence,* pp.142-149. Morgan Kaufmann.

Ezawa, K., J. 1994. Value of Evidence on Influence Diagrams. In R. L. de Mantaras and D. Poole (eds.), *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*, pp. 212-220, Morgan Kaufmann.

Gorry, G. A. 1973. Computer-assisted clinical decision making. *Computer-assisted clinical decision making*, **12**:45-51.

Heckerman, D., Horvitz E., and Middleton B. 1991. An Approximate Non-myopic Computation for Value of Information. In *Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence*, pp. 135-141, Morgan Kauffman.

Heckerman, D. E., Horvitz, E. J. and Nathwani, B. N. 1992. Toward Normative Expert Systems: Part I The Pathfinder Project. *Methods of Information in Medicine*, **31**:90-105, 1992.

Heckerman, D., & Shachter, R. 1995. Decision-Theoretic Foundations for Causal Reasoning. *Journal of Artificial Intelligence Research*, **3**: 405-430.

Howard, R.A. (1960) Dynamic Programming and Markov Processes, the MIT press.

Howard, R.A. 1966. Information value theory. *IEEE Transactions on Systems Science and Cybernetics*, SSC-**2**(1):22-26.

Howard, R.A. 1967. Value of information lotteries. *IEEE Transactions on Systems Science and Cybernetics*, SSC-**3**(1), 54-60.

Howard, R. A. & Metheson, J. E. 1981, Influence Diagrams. In R.A. Howard and J.E. Metheson (eds.), *Readings on the principles and applications of decision analysis*, Vol. 2, pp 719-762, SDG, 1984.

Jensen, F., Jensen, F. V., and Dittmer, S., L. 1994. From Influence Diagrams to Junction Trees. In R. L. de Mantaras and D. Poole (eds.), *Proceedings of the Tenth Conference on Uncertainty in Artificial Intelligence*, pp.367-373, Morgan Kaufmann.

Jensen, F. V., Lauritzen, S. L., and Olesen, K.G. 1990. Bayesian Updating in Causal Probabilistic Networks by Local Computations. *Computational Satistics Quarterly*, (4):269-282.

Kjærulff, U. 1992. A Computational Scheme for Reasoning in Dynamic Probabilistic Networks. In *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence*, pp. 121-129.

Kjærulff, U. 1995. dHugin: a computational system for dynamic time-sliced Bayesian networks. *International Journal of Forecasting*, 11: 89-111.

Pearl, J 1988. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann.

Pearl, J., Geiger, D. & Verma, T. 1990. The logic of influence diagram. In Oliver, R.M. & Smith, J.Q. (eds.), *Influence diagrams, belief nets and decision analysis*, pp 67-87, John Wiley.

Petitti, D. 1994. *Meta-Analysis Decision Analysis and Cost-Effectiveness Analysis*, Oxford University Press, Inc.

Poh, K.L., & Horvitz, E. 1993. Reasoning about the Value of Decision Model Refinement: Methods and Application. In *Proceedings of Ninth Conference on Uncertainty in Artificial Intelligence*, pp.174-182, Morgan Kaufmann.

Poh, K.L., & Horvitz, E. 1996. A Graph-Theoretic Analysis of Information Value. In E. Horvitz and F. Jensen (editors), *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence,* pp. 427-435, Morgan Kaufmann.

Raiffa, H. 1968. *Decision Analysis, Introductory Lectures on Choices under Uncertainty*. Random House.

Shachter, R. D. 1986. Evaluating Influence Diagrams. *Operations Research*, **34**:871-882.

Shachter, R. D. 1990. Evidence Absorption and Propagation through Evidence Reversals. In *Uncertainty in Artificial Intelligence* 5, pp. 173-190. North-Holland.

Shachter, R. D. 1998. Bayes-Ball: The Rational Pastime (for Determining Irrelevance and Requisite Information in Belief Networks and Influence Diagrams), In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence,* pp.480-487, Morgan Kaufmann.

Shachter, R. D. 1999. Efficient Value of Information Computation. In *Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence,* pp.594-601, Morgan Kaufmann.

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

Verma, T., & Pearl, J. 1990. Causal Networks: Semantics and Expressiveness. In *Uncertainty in Artificial Intelligence 4*, pp. 69-7, North Holland.

Xiang, Y. 1999. Temporally Invariant Junction Tree for Inference in Dynamic Bayesian Network, Invited contribution in *Artificial Intelligence Today: Recent Trends and Developments*, pp 473-487, Springer.

Xu, S., Poh, K.L., and Horvitz, E.,J. 2001. Graphical Analysis of Value of Information in Decision Models, in *Proceedings of Fourteenth International Florida Artificial Intelligence Research Society Conference,* 155-159.

Xu, S., and Poh, K.L., 2002. Value of Information Analysis in Dynamic Influence Diagram. A. Abraham, M. Koeppen (Eds.), *Hybrid Information Systems*, pp.597-610, Physica Verlag, Heidelberg.

Norsys Software Corp. 1998. Netica, *Application for Belief Networks and Influence Diagrams.*
(http: //www.norsys.com)

Decision Systems Laboratory. University of Pittsburgh 1998. GeNIe and SMILE,
(http: //www2.sis.pitt.edu/~genie/)