SOLVING INFLUENCE DIAGRAMS USING GIBBS SAMPLING

Ali Jenzarli †

College of Business, The University of Tampa, Plant Hall, Tampa, FL 33606-1490, USA.
Tel: (813) 253-3333 ext. 3306, Fax: (813) 258-7408, Internet: ajenzarl@cfrvm.cfr.usf.edu

ABSTRACT

We describe a Monte Carlo method for solving influence diagrams. This method is a combination of stochastic dynamic programming and Gibbs sampling, an iterative Markov chain Monte Carlo algorithm. Our method is especially useful when exact methods for solving influence diagrams fail.

1. INTRODUCTION

In this paper we describe solution algorithms for influence diagrams (IDs) as elaborations of various forms of the principle of optimality in stochastic dynamic programming, which allows us to find the decision functions in problems of this type sequentially (Bellman and Dreyfus, 1962). In particular, we describe the standard version of the principle of optimality in stochastic dynamic programming. However, since this version is not quite adequate for the case of influence diagrams, we introduce a modification that allows us to determine an optimal decision function for each decision variable sequentially.

In its standard form, the principle of optimality in stochastic dynamic programming applies when we want to maximize or minimize the expectation of a real-valued variable V whose joint distribution with k+1 other variables $\Gamma_0, \Gamma_1, \ldots, \Gamma_k$ (which may each actually be vectors of variables) depends in a stagewise manner on k parameters (which also may be single numbers, vectors or functions) $\delta_1, \ldots, \delta_k$. More precisely, we assume that we can factor the joint probability for $\Gamma_0, \Gamma_1, \ldots, \Gamma_k$ and V in the form

$$P_{\delta_1,\ldots,\delta_k}(\Gamma_0,\ldots,\Gamma_k,V) = h_0(\Gamma_0)h_{\delta_1}(\Gamma_1|\Gamma_0)\ldots h_{\delta_{k-1}}(\Gamma_{k-1}|\Gamma_0,\ldots,\Gamma_{k-2})h_{\delta_k}(\Gamma_k,V|\Gamma_0,\ldots,\Gamma_{k-1}), \quad (1)$$

where the factors are conditional probabilities. We must also assume that it is computationally feasible to compute $E_{\delta_k}(V|\Gamma_0,\ldots,\Gamma_{k-1})$ from $h_{\delta_k}(\Gamma_k,V|\Gamma_0,\ldots,\Gamma_{k-1})$ for each value of δ_k and each configuration of values of $\Gamma_0 \cup \ldots \cup \Gamma_{k-1}$, or at least to find for each configuration $(\gamma_0,\ldots,\gamma_{k-1})$ of $\Gamma_0 \cup \ldots \cup \Gamma_{k-1}$ the value of δ_k that optimizes

$$E_{\delta_k}(V|\Gamma_0 = \gamma_0,\ldots,\Gamma_{k-1} = \gamma_{k-1}). \quad (2)$$

Finally, we must assume (this is crucial) that we can find a single value of δ_k that optimizes (2) for <u>all</u> $(\gamma_0,\ldots,\gamma_{k-1})$. Since the distribution of $\Gamma_0 \cup \ldots \cup \Gamma_{k-1}$ does not depend on δ_k, we have

$$E_{\delta_1,\ldots,\delta_k}(V) = E_{\delta_1,\ldots,\delta_{k-1}}\left(E_{\delta_k}(V|\Gamma_0,\ldots,\Gamma_{k-1})\right).$$

Therefore, this optimizing value of δ_k will also optimize the unconditional expectation $E_{\delta_1,\ldots,\delta_k}(V)$ for any choice of $(\delta_1,\ldots,\delta_{k-1})$. And therefore, it can be extended to a choice of $(\delta_1,\ldots,\delta_k)$ to optimize this unconditional expectation.

† The author is currently Assistant Professor of Management at the University of Tampa. He is finishing his doctoral dissertation at the University of Kansas under professor Glenn Shafer (Rutgers University and Princeton University), and professor Prakash Shenoy (The University of Kansas).

Suppose we fix this optimal value of δ_k eliminating it from our notation, and reducing (1) to

$$h_{\delta_1,\dots,\delta_{k-1}}(\Gamma_0,\dots,\Gamma_k,V)=h_0(\Gamma_0)h_{\delta_1}(\Gamma_1|\Gamma_0)\dots h_{\delta_{k-1}}(\Gamma_{k-1}|\Gamma_0,\dots,\Gamma_{k-2})h(\Gamma_k,V|\Gamma_0,\dots,\Gamma_{k-1}), \quad (3)$$

From this point, we proceed in either of two ways. We can sum or integrate Γ_k out of the expectation. Or we can incorporate Γ_k as part of Γ_{k-1}.

The first option, summing or integrating Γ_k out, means reducing (3) to $h_{\delta_1,\dots,\delta_{k-1}}(\Gamma_0,\dots,\Gamma_{k-1},V)=h_0(\Gamma_0)h_{\delta_1}(\Gamma_1|\Gamma_0)\dots h'_{\delta_{k-1}}(\Gamma_{k-1},V|\Gamma_0,\dots,\Gamma_{k-2})$, where $h'_{\delta_{k-1}}(\Gamma_{k-1},V|\Gamma_0,\dots,\Gamma_{k-2})=h_{\delta_{k-1}}(\Gamma_{k-1}|\Gamma_0,\dots,\Gamma_{k-2})\int h(\gamma_k,V|\Gamma_0,\dots,\Gamma_{k-1})d\gamma_k$. Once again, we assume that we can choose δ_{k-1} so as to optimize simultaneously

$$E_{\delta_{k-1}}(V|\Gamma_0=\gamma_0,\Gamma_1=\gamma_1,\dots,\Gamma_{k-2}=\gamma_{k-2}) \quad (4)$$

for all $(\gamma_0,\dots,\gamma_{k-2})$. Then, as before, the choice of δ_{k-1} can be extended to a choice of $(\delta_1,\dots,\delta_{k-1})$ to optimize the unconditional expectation $E_{\delta_1,\dots,\delta_{k-1}}(V)$.

So we may also fix this optimal value of δ_{k-1}, and reduce the problem further. We can continue in this way, choosing the δ_i sequentially, provided that the successive simultaneous optimizations like those in (2), (4), etc. are possible.

The second option means setting $\Gamma'_{k-1}=\Gamma_{k-1}\cup\Gamma_k$, and reducing (3) to $h_{\delta_1,\dots,\delta_{k-1}}(\Gamma_0,\dots,\Gamma_{k-1},V)=h_0(\Gamma_0)h_{\delta_1}(\Gamma_1|\Gamma_0)\dots h'_{\delta_{k-1}}(\Gamma'_{k-1},V|\Gamma_0,\dots,\Gamma_{k-2})$, where $h'_{\delta_{k-1}}(\Gamma'_{k-1},V|\Gamma_0,\dots,\Gamma_{k-2})=h_{\delta_{k-1}}(\Gamma_{k-1}|\Gamma_0,\dots,\Gamma_{k-2})h(\Gamma_k,V|\Gamma_0,\dots,\Gamma_{k-1})$. Again, if we can choose δ_{k-1} to optimize simultaneously (4) for all $(\gamma_0,\dots,\gamma_{k-2})$, and so on, we can proceed to choose the δ_i sequentially.

This standard version of stochastic dynamic programming is not quite adequate for the case of influence diagrams. The reason is that although these diagrams involve factorizations that can be written in the form (1), the factors are not necessarily conditional probabilities.

The standard version of stochastic dynamic programming can be modified to fit influence diagrams, but there has been a considerable variety of opinion about how to do this. The oldest sequential solution algorithm for influence diagrams, the Olmsted-Shachter reduction algorithm (Olmsted, 1983; Shachter, 1986) goes considerably beyond stochastic dynamic programming, in order to maintain a representation of the influence diagram form as the algorithm proceeds. More recent algorithms, including the valuation network algorithm of Shenoy (1992 and 1993) and the potential influence diagram algorithm of Ndilikilikesha (1992), stay closer to stochastic dynamic programming.

The simulation algorithm we describe in Section 2 does not fit exactly into either Shenoy's or Ndilikilikesha's framework, primarily because their algorithms integrate Γ_k out, while our algorithm follows the second option described above, that of absorbing Γ_k into $h_{\delta_{k-1}}$. We could elaborate one of their frameworks in order to make our algorithm fit, but it will be simpler for us to deal directly with the necessary modification in the standard form of stochastic dynamic programming that we have just described.

Here is the modification that we require. Let us assume that the joint probability for $\Gamma_0,\Gamma_1,\dots,\Gamma_k$ and V is proportional to a factorization of the following form

$$P_{\delta_1,\ldots,\delta_k}(\Gamma_0,\ldots,\Gamma_k,V) \propto h_0(\Gamma_0)h_{\delta_1(\Gamma_0)}(\Gamma_1|\Gamma_0)\ldots h_{\delta_{k-1}(\Gamma_0,\ldots,\Gamma_{k-2})}(\Gamma_{k-1}|\Gamma_0,\ldots,\Gamma_{k-2})h_{\delta_k(\Gamma_0,\ldots,\Gamma_{k-1})}(\Gamma_k,V|\Gamma_0,\ldots,\Gamma_{k-1}) \quad (5)$$

Here we do <u>not</u> assume that the factors are conditional probabilities. But we <u>do</u> assume that the δ_i are functions; and we assume, as the notation indicates, that for fixed values of $\Gamma_0,\Gamma_1,\ldots,\Gamma_{k-1}$, the factor $h_{\delta_k(\Gamma_0,\ldots,\Gamma_{k-1})}$, regarded as a function of Γ_k and V, depends on δ_k only through the value δ_k assigns to those values of $\Gamma_0,\Gamma_1,\ldots,\Gamma_{k-1}$. This assumption, as we will see implies that the simultaneous optimizations at each step are possible.

Notice first that the factorization (5) implies that $h_{\delta_k(\Gamma_0,\ldots,\Gamma_{k-1})}(\Gamma_k,V|\Gamma_0,\ldots,\Gamma_{k-1})$, for fixed values of $\Gamma_0,\Gamma_1,\ldots,\Gamma_{k-1}$, is at least proportional to the conditional probability distribution for Γ_k and V given these values of $\Gamma_0,\Gamma_1,\ldots,\Gamma_{k-1}$. To see this, recall that a conditional probability distribution is always proportional to the corresponding joint probability distribution. Thus

$$P_{\delta_1,\ldots,\delta_k}(\Gamma_k,V|\Gamma_0,\ldots,\Gamma_{k-1}) = \lambda P_{\delta_1,\ldots,\delta_k}(\Gamma_0,\ldots\Gamma_k,V), \quad (6)$$

where λ is constant with respect to Γ_k and V. (The other variables are thought of as fixed.) We usually write (6) with a symbol of proportionality:

$$P_{\delta_1,\ldots,\delta_k}(\Gamma_k,V|\Gamma_0,\ldots,\Gamma_{k-1}) \propto P_{\delta_1,\ldots,\delta_k}(\Gamma_0,\ldots\Gamma_k,V).$$

Since only the last factor of (5) involves Γ_k or V, (6) implies that

$$P_{\delta_1,\ldots,\delta_k}(\Gamma_k,V|\Gamma_0,\ldots,\Gamma_{k-1}) \propto h_{\delta_k(\Gamma_0,\ldots,\Gamma_{k-1})}(\Gamma_k,V|\Gamma_0,\ldots,\Gamma_{k-1}).$$

Again, this proportionality is to be interpreted by taking both sides as functions of Γ_k and V only, with the other variables fixed; we are able to omit the other factors only because they, as functions of the other variables, are also fixed and hence can be absorbed into the constant of proportionality.

Whenever a function is proportional to a probability distribution (or probability conditional), it contains all the information needed to find that conditional because the constant of proportionality is simply what is needed to make the function sum (or integrate) to one. Thus

$$h_{\delta_k(\gamma_0,\ldots,\gamma_{k-1})}(\Gamma_k,V|\Gamma_0=\gamma_0,\ldots,\Gamma_{k-1}=\gamma_{k-1})$$

has, in particular, all the information needed to determine the conditional expectation of V given $(\gamma_0,\ldots,\gamma_{k-1})$,

$$E_{\delta_k(\gamma_0,\ldots,\gamma_{k-1})}(\Gamma_k,V|\Gamma_0=\gamma_0,\ldots,\Gamma_{k-1}=\gamma_{k-1}). \quad (7)$$

We can choose the value of $\delta_k(\gamma_0,\ldots,\gamma_{k-1})$ to optimize this expectation, and by doing this for each set of values $(\gamma_0,\ldots,\gamma_{k-1})$, we will have chosen a function δ_k that simultaneously optimizes (7) for all $(\gamma_0,\ldots,\gamma_{k-1})$.

Once this choice of δ_k has been carried out, we can proceed, as before, absorbing $h_{\delta_k(\Gamma_0,\ldots,\Gamma_{k-1})}$ into $h_{\delta_{k-1}(\Gamma_0,\ldots,\Gamma_{k-2})}$, first integrating γ_k out if we wish to do so.

In order to fit influence diagrams into this version of stochastic dynamic programming, we write Γ_i for the set of variables consisting of Δ_i together with the chance variables observed by the decision maker between Δ_i and Δ_{i+1}, for i = 1, ..., k-1, we write Γ_0 for the chance variables observed before Δ_1 and Γ_k for the set of variables consisting of Δ_k together with the chance variables (other than V) observed after Δ_k (or never), and we write δ_i for the decision function for Δ_i. Then we set $h_0(\Gamma_0)$ equal to the product of conditionals for the chance variables in Γ_0. For i = 1, ..., k-1, we set $h_{\delta_i}(\Gamma_i|\Gamma_0,\ldots,\Gamma_{i-1})$ equal to the product of conditionals for the chance variables in Γ_i, times the conditional corresponding to the decision function δ_i (this conditional gives only probabilities of zero and one). And we similarly set $h_{\delta_k}(\Gamma_k, V|\Gamma_0,...,\Gamma_{k-1})$ equal to the product of the conditionals for all variables in $\Gamma_k \cup \{V\}$. Since $h_{\delta_k}(\Gamma_k, V|\Gamma_0,...,\Gamma_{k-1})$ depends on δ_k only through its value $\delta_k(\Gamma_0,...,\Gamma_{k-1})$, this puts us in the framework just described.

2. DERIVATION OF A GENERAL SOLUTION ALGORITHM

In this section we show how to use Gibbs sampling (Geman and Geman, 1984; and Gelfand and Smith, 1990), an iterative Markov chain Monte Carlo algorithm (Hastings,) to implement stochastic dynamic programming for an influence diagram. Since the stochastic dynamic program is iterative, it suffices to explain how to implement it using Gibbs sampling for a single step. We will explain how to implement it for the first step.

Our task, then, is to find the decision function δ_k. This means finding, for each configuration $(\gamma_0, \gamma_1,...,\gamma_{k-1})$ of $\Gamma_0 \cup ... \cup \Gamma_{k-1}$, the value d_k of the decision Δ_k that optimizes

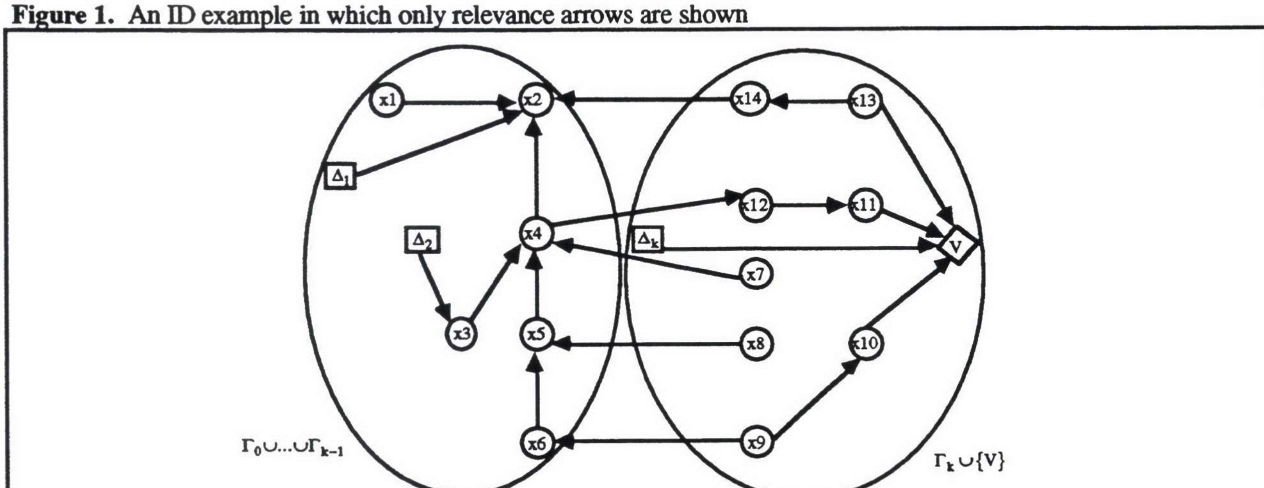
$$E_{d_k}(V|\Gamma_0 = \gamma_0, \Gamma_1 = \gamma_1,...,\Gamma_{k-1} = \gamma_{k-1}). \tag{8}$$

(Notice that we write d_k in the place of δ_k as a subscript on the expectation operator; this is because the expectation for the configuration $(\gamma_0, \gamma_1,...,\gamma_{k-1})$ of the predecessors depends only on the value d_k that δ_k assigns to this configuration.) To this end, we simply compute (8) for all d_k and choose the d_k that gives the optimal (largest or smallest depending on whether we are maximizing or minimizing) result.

To compute (8) for a particular d_k, we recall that the conditional joint distribution of $\Gamma_k \cup \{V\}$ is proportional to $h_{d_k}(\Gamma_k, V|\Gamma_0 = \gamma_0,...,\Gamma_{k-1} = \gamma_{k-1})$, which is simply the product of the conditionals for $\Gamma_k \cup \{V\}$.

Leaving aside the variables Δ_k and V, which are deterministic in this conditional joint distribution (Δ_k is equal to the constant d_k, and V is a function of the other variables), we can say that the conditional distribution of the other variables (all chance variables) is the product of their original conditionals. We are not interested, however, in all these variables; we are really interested only in V. Hence we can discard the conditionals for any variables that are independent of V in the conditional joint distribution. Figure 1 gives an example where relevance arrows are shown, and informational arrows are omitted (Clemens, 1991).

Figure 1. An ID example in which only relevance arrows are shown



Here x7 and x8 are independent of V in the conditional distribution of $\Gamma_k \cup \{V\}$ because the distribution factors into parts involving only x7 and x8 and parts involving only the other variables. So we may omit their conditionals, effectively eliminating them from the problem, and reducing Γ_k to a smaller set Γ_k'.

The general rule for this reduction of our problem can be formulated graphically as follows. Consider the directed graph of the ID without informational arrows (as in Figure 1). Form the moral graph (Jensen et al., 1990). And omit any variables from Γ_k that are not connected with V in the subgraph of this moral graph determined by $\Gamma_k \cup \{V\}$. A variable X in Γ_k is not connected with V in this subgraph if there is no path in $\Gamma_k \cup \{V\}$ that connects X to V. Figures 2 and 3 show how this procedure applies to the example of Figure 1.

Figure 2. Moral graph with x7 and x8

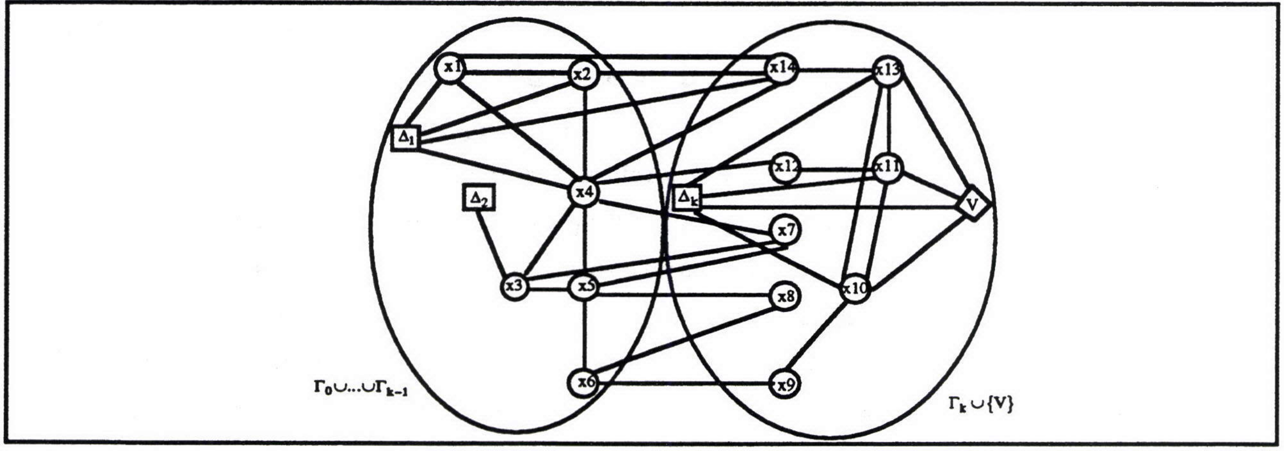
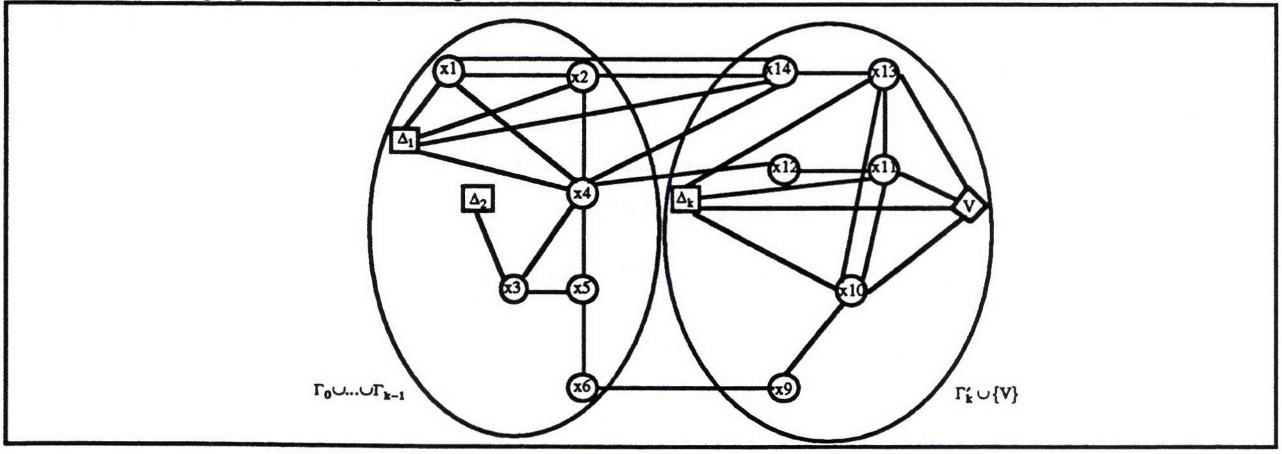


Figure 3. Moral graph without x7 and x8



Next, notice that only some of the variables in $\Gamma_0 \cup ... \cup \Gamma_{k-1}$ will affect the expectation (8). Indeed, the only ones to affect it will be those involved in the factors that remain in the product. In our example, these factors are $\Gamma_k' \cup \{V\}$

1. $f_1(x_{14}) = [x_2 = \bullet | x_1 = \bullet, x_4 = \bullet, x_{14}, \Delta_1 = \bullet]$
2. $f_2(x_{13}, x_{14}) = [x_{14} | x_{13}]$

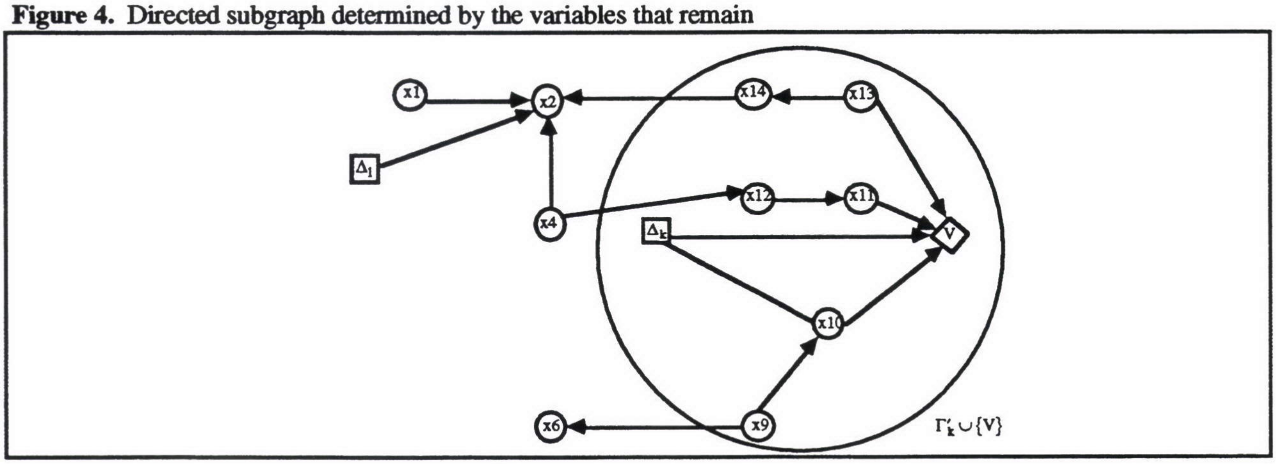
3. $f_3(x_{12}) = [x_{12}|x_4 = \bullet]$
4. $f_4(x_{11}, x_{12}) = [x_{11}|x_{12}]$
5. $f_5(x_9, x_{10}) = [x_{10}|x_9]$
6. $f_6(x_9) = [x_6 = \bullet|x_9]$
7. $f_7(x_{13}) = [x_{13}]$
8. $f_8(x_9) = [x_9]$

The variables involved in these factors are x_1, x_2, x_4, x_6 and Δ_1. The other variables, Δ_2, x_3 and x_5, are not involved. Since the expectation (8) is not affected by the variables in $\Gamma_0 \cup \ldots \cup \Gamma_{k-1}$ that are not involved in these factors, we need not make our choice of d_k depend on them. In other words, we can make Δ_k a function only of the variables involved in the factors—x_1, x_2, x_4, x_6 and Δ_1 in our example.

The general rule for finding the variables on which Δ_k will depend can be described in terms of the moral graph we obtained previously: they are the neighbors of Γ_k in this graph. In our example (Figure 3), we see that x_1, x_2, x_4, x_6 and Δ_1 are the neighbors of Γ_k.

The factors that remain can be envisioned in terms of the directed subgraph determined by the variables that remain, as in Figure 4. They are the factors in which the variables remaining in the circle are parents or children.

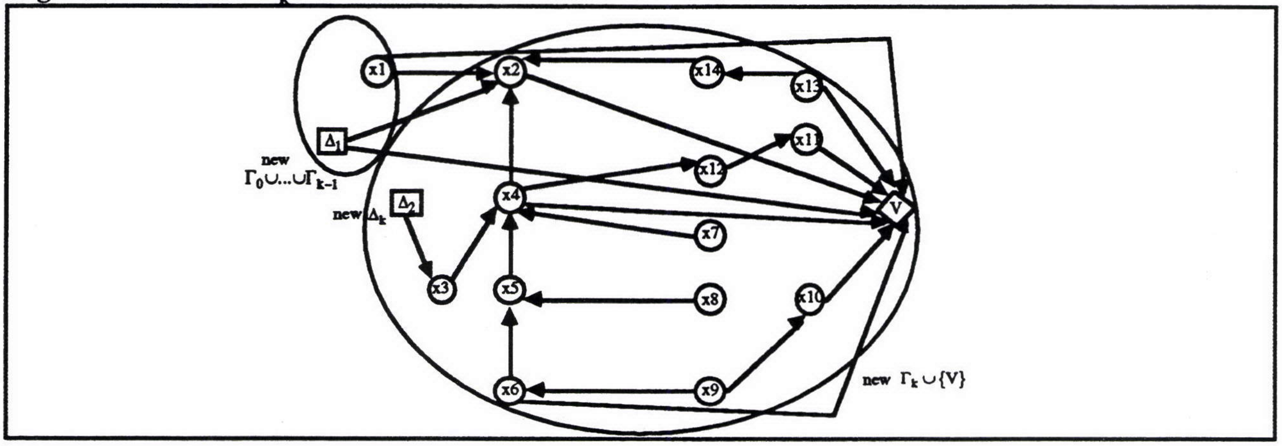
Figure 4. Directed subgraph determined by the variables that remain



Now that the factors are identified, we do Gibbs sampling with these factors to simulate the joint distribution of $\Gamma'_k \setminus \{\Delta_k, V\}$. For the configuration of $\Gamma'_k \setminus \{\Delta_k, V\}$ obtained at each step of the Gibbs sampling, we compute V. This gives a sequence of values for V simulating a random sample from its conditional distribution, from which we may compute its conditional expectation.

When we move on to the next step of the stochastic dynamic program, we use the second of the two options discussed in Section 1. In other words, we absorb Γ_k into Γ_{k-1}, and we include the conditionals from Γ_k in the new factorization of $h_{\delta_{k-1}}$. In order to avoid zero probabilities that would interfere with the Gibbs sampling, we do not include the conditional for Δ_k corresponding to the decision function we have just found for Δ_k. Instead, we substitute this decision function in all the conditionals in which Δ_k appeared as a parent, thus eliminating Δ_k from the graph and producing arrows from the variables on which Δ_k depends to the variables for which it was a parent. Figure 5 illustrates the result for our example.

Figure 5. DAG with Δ_k absorbed into its direct successors



Acknowledgments

This work was partially supported by the National Science Foundation under grants IRI–8902444 and SES–9213558, and by a grant from the General Research Fund at the University of Kansas, grant 3605–XX–0038. I am grateful for comments from my mentors Glenn Shafer and Prakash Shenoy. I am especially grateful for the help and encouragement I received from Glenn Shafer.

References

Bellman, R. and Dreyfus, S. (1962) *Applied Dynamic Programming*, Princeton University Press, Princeton, N.J.

Gelfand, A. E. and Smith, A. F. M. (1990) Sampling-based approaches to calculating marginal densities, *Journal of the American Statistical Association*, **85**(410), 398-409.

Geman, S. and Geman, D. (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAMI-6**, 721-741.

Hastings, W. K. (1970) Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, **57**(1), 97-109.

Jensen, F. V., Lauritzen, S. L., and Olesen, K. G. (1990) Bayesian updating in causal probabilistic networks by local computations, *Computational Statistics Quarterly* **4**, 269-282.

Ndilikilikesha, P. (1992) Potential influence diagrams, *Working Paper No. 235*, School of Business, University of Kansas, Lawrence, Kansas.

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

Shachter, R. D. (1986) Evaluating influence diagrams, *Operations Research*, **34**(6), 871-882.

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

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