Goal Oriented Symbolic Propagation in Bayesian Networks

Enrique Castillo*, **José Manuel Gutiérrez*** and **Ali S. Hadi****
* Department of Applied Mathematics and Computational Sciences,
University of Cantabria, SPAIN
Castie@ccaix3.unican.es and Gutierjm@ccaix3.unican.es
** Department of Social Statistics, Cornell University, USA
Ali-hadi@cornell.edu

Abstract

The paper presents an efficient goal oriented algorithm for symbolic propagation in Bayesian networks. The proposed algorithm performs symbolic propagation using numerical methods. It first takes advantage of the independence relationships among the variables and produce a reduced graph which contains only the relevant nodes and parameters required to compute the desired propagation. Then, the symbolic expression of the solution is obtained by performing numerical propagations associated with specific values of the symbolic parameters. These specific values are called the canonical components. Substantial savings are obtained with this new algorithm. Furthermore, the canonical components allow us to obtain lower and upper bounds for the symbolic expressions resulting from the propagation. An example is used to illustrate the proposed methodology.

Introduction

Bayesian networks are powerful tools both for graphically representing the relationships among a set of variables and for dealing with uncertainties in expert systems. A key problem in Bayesian networks is evidence propagation, that is, obtaining the posterior distributions of the variables when some evidence is observed. Several efficient exact and approximate methods for propagation of evidence in Bayesian networks have been proposed in recent years (see, for example, Pearl 1988, Lauritzen and Spiegelhalter 1988, Henrion 1988, Shachter and Peot 1990, Fung and Chang 1990, Poole 1993, Bouckaert, Castillo and Gutiérrez 1995). However, these methods require that the joint probabilities of the nodes be specified numerically, that is, all the parameters must be assigned numeric values. In practice, when exact numeric specification of these parameters may not be available, or sensitivity analysis is desired, there is a need for symbolic methods which are able to deal with the parameters themselves, without assigning them numeric values. Symbolic propagation leads to solutions which are expressed as functions of the parameters in symbolic form.

Recently, two main approaches have been proposed for symbolic inference in Bayesian networks. The symbolic probabilistic inference algorithm (SPI) (Shachter, D'Ambrosio and DelFabero 1990 and Li and D'Ambrosio 1994) is a goal oriented method which performs only those calculations that are required to respond to queries. Symbolic expressions can be obtained by postponing evaluation of expressions, maintaining them in symbolic form. On the other hand, Castillo, Gutiérrez and Hadi 1995, 1996a, 1996b, exploit the polynomial structure of the marginal and conditional probabilities in Bayesian networks to efficiently perform symbolic propagation by calculating the associated numerical coefficients using any standard numeric method for inference in Bayesian networks. As opposed to the SPI algorithm, this method is not goal oriented, but allows us to obtain symbolic expressions for all the nodes in the network. In this paper we show that this algorithm is also suitable for goal oriented problems. In this case, the performance of the method can be improved by taking advantage of the independence relationships among the variables and produce a reduced graph which contains only the nodes relevant to the desired propagation. Thus, only those operations required to obtain the desired computations are performed.

We start by introducing the necessary notation. Then, an algorithm for efficient computation of the desired conditional probabilities is presented and its practical application is illustrated by an example. Finally, we show how to obtain lower and upper bounds for the symbolic expressions solution of the given problem.

Notation

Let $X = \{X_1, X_2, \ldots, X_n\}$ be a set of n discrete variables, each can take values in the set $\{0, 1, \ldots, r_i\}$, the possible states of the variable X_i. A Bayesian network over X is a pair (D, P), where the graph D is a directed acyclic graph (DAG) with one node for each variable in X and $P = \{P_1(x_1|\pi_1), \ldots, P_n(x_n|\pi_n)\}$ is a set of n conditional probabilities, one for each variable, where Π_i is the set of parents of node X_i. Using the chain rule, the joint probability density of X can be written as:

$$P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P_i(x_i|\pi_i). \tag{1}$$

Some of the conditional probability distributions in (1) can be specified numerically and others symbolically, that is, $P_i(x_i|\pi_i)$ can be a parametric family. When $P_i(x_i|\pi_i)$ is a parametric family, we refer to the node X_i as a symbolic node. A convenient notation for the parameters in this case is given by

$$\theta_{ij\pi} = P_i(X_i = j|\Pi_i = \pi),\ j \in \{0, \ldots, r_i\}, \tag{2}$$

where π is any possible instantiation of the parents of X_i. Thus, the first subscript in $\theta_{ij\pi}$ refers to the node number, the second subscript refers to the state of the node, and the remaining subscripts refer to the parents' instantiations. Since $\sum_{j=0}^{r_i} \theta_{ij\pi} = 1$, for all i and π, any one of the parameters can be written as one minus the sum of all others. For example, $\theta_{ir_i\pi}$ is

$$\theta_{ir_i\pi} = 1 - \sum_{j=0}^{r_i-1} \theta_{ij\pi}. \tag{3}$$

If X_i has no parents, we use θ_{ij} to denote $P_i(X_i = j)$, $j \in \{0, \ldots, r_i\}$, for simplicity.

Goal Oriented Algorithm

Suppose that we are interested in a given goal node X_i, and that we want to obtain the conditional probabilities $P(X_i = j|E = e)$, where E is a set of evidential nodes with known values $E = e$. Using the algebraic characterization of the probabilities given by Castillo, Gutiérrez and Hadi 1995, the unnormalized probabilities $\hat{P}(X_i = j|E = e)$ are polynomials of the form:

$$\hat{P}(X_i = j|E = e) = \sum_{m_r \in M_j} c_{jr} m_r = p_j(\Theta), \tag{4}$$

where m_j are monomials in the symbolic parameters, Θ, contained in the probability distribution of the Bayesian network.

For example, suppose we have a discrete Bayesian network consisting of five binary variables $\{X_1, \ldots, X_5\}$, with values in the set $\{0, 1\}$. The associated DAG is given in Figure 1. Table 1 gives the corresponding parameters, some in numeric and others in symbolic form. Node X_4 is numeric because it contains only numeric parameters and the other four nodes are symbolic because some of their parameters are specified only symbolically.

For illustrative purposes, suppose that the target node is X_3 and that we have the evidence $X_2 = 1$. We wish to compute the conditional probabilities $P(X_3 = j|X_2 = 1), j = 0, 1$. We shall show that

$$P(X_3 = 0|X_2 = 1) = \frac{0.4\theta_{10}\theta_{210} + 0.3\theta_{301} - 0.3\theta_{10}\theta_{301}}{0.3 - 0.3\theta_{10} + \theta_{10}\theta_{210}} \tag{5}$$

Node X_i	Node Π_i	Parameters $X_i = 0$
X_1	None	$\theta_{10} = P(X_1 = 0)$
X_2	X_1	$\theta_{200} = P(X_2 = 0\|X_1 = 0)$ $\theta_{201} = P(X_2 = 0\|X_1 = 1) = 0.7$
X_3	X_1	$\theta_{300} = P(X_3 = 0\|X_1 = 0) = 0.4$ $\theta_{301} = P(X_3 = 0\|X_1 = 1)$
X_4	X_2, X_3	$\theta_{4000} = P(X_4 = 0\|X_2 = 0, X_3 = 0) = 0.2$ $\theta_{4001} = P(X_4 = 0\|X_2 = 0, X_3 = 1) = 0.4$ $\theta_{4010} = P(X_4 = 0\|X_2 = 1, X_3 = 0) = 0.7$ $\theta_{4011} = P(X_4 = 0\|X_2 = 1, X_3 = 1) = 0.8$
X_5	X_3	$\theta_{500} = P(X_5 = 0\|X_3 = 0)$ $\theta_{501} = P(X_5 = 0\|X_3 = 1)$

Node X_i	Node Π_i	Parameters $X_i = 1$
X_1	None	$\theta_{11} = P(X_1 = 1)$
X_2	X_1	$\theta_{210} = P(X_2 = 1\|X_1 = 0)$ $\theta_{211} = P(X_2 = 1\|X_1 = 1) = 0.3$
X_3	X_1	$\theta_{310} = P(X_3 = 1\|X_1 = 0) = 0.6$ $\theta_{311} = P(X_3 = 1\|X_1 = 1)$
X_4	X_2, X_3	$\theta_{4100} = P(X_4 = 1\|X_2 = 0, X_3 = 0) = 0.8$ $\theta_{4101} = P(X_4 = 1\|X_2 = 0, X_3 = 1) = 0.6$ $\theta_{4110} = P(X_4 = 1\|X_2 = 1, X_3 = 0) = 0.3$ $\theta_{4111} = P(X_4 = 1\|X_2 = 1, X_3 = 1) = 0.2$
X_5	X_3	$\theta_{510} = P(X_5 = 1\|X_3 = 0)$ $\theta_{511} = P(X_5 = 1\|X_3 = 1)$

Table 1: Numeric and symbolic conditional probabilities.

and

$$P(X_3 = 1|X_2 = 1) = \frac{0.3 - 0.3\theta_{10} + 0.6\theta_{10}\theta_{210} - 0.3\theta_{301} + 0.3\theta_{10}\theta_{301}}{0.3 - 0.3\theta_{10} + \theta_{10}\theta_{210}}. \tag{6}$$

where the denominators are the corresponding normalizing constants.

Algorithm 1 gives the solution for this goal oriented problem by calculating the coefficients c_{jr} in (4) of these polynomials. It is organized in four main parts:

- **PART I : Identify all Relevant Nodes.**
  The conditional probability $P(X_i = j|E = e)$ does

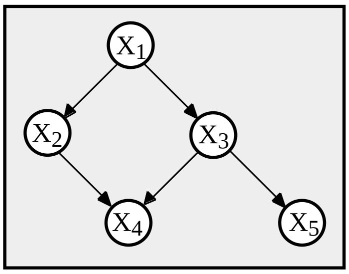

Figure 1: An example of a five-node Bayesian Network.

not necessarily involve parameters associated with all nodes. Thus, we identify the set of nodes which are relevant to the calculation of $P(X_i = j|E = e)$, using either one of the two algorithms given in Geiger, Verma, and Pearl 1990 and Shachter 1990. Once this has been done we can remove the remaining nodes from the graph and identify the associated set of relevant parameters Θ.

- **PART II : Identify Sufficient Parameters.** By considering the values of the evidence variables, the set of parameters Θ can be further reduced by identifying and eliminating the set of parameters which are in contradiction with the evidence. These parameters are eliminated using the following two rules:
  - **Rule 1:** Eliminate the parameters $\theta_{jk\pi}$ if $x_j \neq k$ for every $X_j \in E$.
  - **Rule 2:** Eliminate the parameters $\theta_{jk\pi}$ if parents' instantiations π are incompatible with the evidence.
- **PART III : Identify Feasible Monomials.** Once the minimal sufficient subsets of parameters has been identified, they are combined in monomials by taking the Cartesian product of the minimal sufficient subsets of parameters and eliminating the set of all infeasible combinations of the parameters using:
  - **Rule 3:** Parameters associated with contradictory conditioning instantiations cannot appear in the same monomial.
- **PART IV : Calculate Coefficients of all Polynomials.** This part calculates the coefficients applying numeric network inference methods to the reduced graph obtained in Part I. If the parameters Θ are assigned numerical values, say θ, then $p_j(\theta)$ can be obtained using any numeric network inference method to compute $P(X_i = j|E = e, \Theta = \theta)$. Similarly, the monomials m_r take a numerical value, the product of the parameters involved in m_r. Thus, we have

$$\hat{P}(X_i = j|E = e, \Theta = \theta) = \sum_{m_r \in M_j} c_{jr} m_r = p_j(\theta). \tag{7}$$

Note that in (7) all the monomials m_r, and the unnormalized probability $p_j(\theta)$ are known numbers, and the only unknowns are the coefficients c_{jr}. To compute these coefficients, we need to construct a set of independent equations each of the form in (7). These equations can be obtained using sets of distinct instantiations Θ.

To illustrate the algorithm we use, in parallel, the previous example.

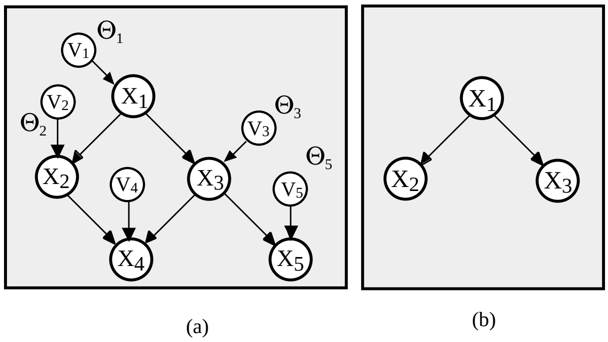

Figure 2: (a) Augmented graph D^* after adding a dummy node V_i for every symbolic node X_i, and (b) the reduced graph D' sufficient to compute $P(X_i = j|E = e)$.

Algorithm 1 Goal Oriented Symbolic Propagation.

Input: A Bayesian network (D, P), a target node X_i and an evidential set E (possibly empty) with evidential values $E = e$.

Output: The probabilities $P(X_i = j|E = e)$.

PART I:

- **Step 1:** Construct a DAG D^* by augmenting D with a dummy node V_j and adding a link $V_j \rightarrow X_j$ for every node X_j in D. The node V_j represents the parameters, Θ_j, of node X_j.
- **Step 1 (example):** We add to the initial graph in Figure 1, the nodes V_1, V_2, V_3, V_4, and V_5 The resulting graph in shown in Figure 2(a).
- **Step 2:** Identify the set V of dummy nodes in D^* not d-separated from the goal node X_i by E. Obtain a new graph D' by removing from D those nodes whose corresponding dummy nodes are not contained in V with the exception of the target and evidential nodes. Let Θ be the set of all the parameters associated with the symbolic nodes included in the new graph and V.
- **Step 2 (example):** The set V of dummy nodes not d-separated from the goal node X_3 by the evidence node $E = \{X_2\}$ is found to be $V = \{V_1, V_2, V_3\}$. Therefore, we remove X_4 and X_5 from the graph obtaining the graph shown in Figure 2(b). Thus, the set of all the parameters associated with symbolic nodes of the new graph is

  $\Theta = \{\Theta_1, \Theta_2, \Theta_3\} = \{\{\theta_{10}, \theta_{11}\};$
  $\{\theta_{200}, \theta_{210}, \theta_{201}, \theta_{211}\}; \{\theta_{300}, \theta_{310}, \theta_{301}, \theta_{311}\}\}.$

PART II:

- **Step 3:** If there is evidence, remove from Θ the parameters $\theta_{jk\pi}$ if $x_j \neq k$ for $X_j \in E$ (Rule 1).

- **Step 3 (example):** The set Θ contains the symbolic parameters θ_{200} and θ_{201} that do not match the evidence $X_2 = 1$. Then, applying Rule 1 we eliminate these parameters from Θ.
- **Step 4:** If there is evidence, remove from Θ the parameters $\theta_{jk\pi}$ if the set of values of parents' instantiations π are incompatible with the evidence (Rule 2).
- **Step 4 (example):** Since the only evidential node X_2 has no sons in the new graph, no further reduction is possible. Thus, we get the minimum set of sufficient parameters:

$$\Theta = \{\{\theta_{10}, \theta_{11}\}; \{\theta_{210}, \theta_{211}\}; \{\theta_{300}, \theta_{310}, \theta_{301}, \theta_{311}\}\}.$$

PART III:

- **Step 5:** Obtain the set of monomials M by taking the Cartesian product of the subsets of parameters in Θ.
- **Step 5 (example):** The initial set of candidate monomials is given by taking the Cartesian product of the minimal sufficient subsets, that is,

$$M = \{\theta_{10}, \theta_{11}\} \times \{\theta_{210}, \theta_{211}\} \times \{\theta_{300}, \theta_{310}, \theta_{301}, \theta_{311}\}.$$

Thus, we obtain 16 different candidate monomials.
- **Step 6:** Using Rule 3, remove from M those monomials which contain a set of incompatible parameters.
- **Step 6 (example):** Some of the monomials in M contain parameters with contradictory instantiations of the parents. For example, the monomial $\theta_{10}\theta_{210}\theta_{301}$ contains contradictory instantiations of the parents because θ_{10} indicates that $X_1 = 0$ whereas θ_{301} indicates that $X_1 = 1$. Thus, applying Rule 3, we get the following set of feasible monomials $M = \{\theta_{10}\theta_{210}\theta_{300}, \theta_{10}\theta_{210}\theta_{310}, \theta_{11}\theta_{211}\theta_{301}, \theta_{11}\theta_{211}\theta_{311}\}$.
- **Step 7:** If some of the parameters associated with the symbolic nodes are specified numerically, then remove these parameters from the resulting feasible monomials because they are part of the numerical coefficients.
- **Step 7 (example):** Some symbolic nodes involve both numeric and symbolic parameters. Then, we remove from the monomials in M the numerical parameters $\theta_{300}, \theta_{310}$ and θ_{211} obtaining the set of feasible monomials $M = \{\theta_{10}\theta_{210}, \theta_{11}\theta_{301}, \theta_{11}\theta_{311}\}$. Note that, when removing these numeric parameters from Θ, the monomials $\theta_{10}\theta_{210}\theta_{300}$ and $\theta_{10}\theta_{210}\theta_{310}$ become $\theta_{10}\theta_{210}$. Thus, finally, we only have three different monomials associated with the probabilities $P(X_3 = j|X_2 = 1), j = 0, 1$.

PART IV:

- **Step 8:** For each possible state j of node X_i, $j = 0, \ldots, r_i - 1$, build the subset M_j by considering those monomials in M which does not contain any parameter of the form $\theta_{iq\pi}$, with $q \neq j$.
- **Step 8 (example):** The sets of monomials needed to calculate $P(X_3 = 0|X_2 = 1)$ and $P(X_3 = 1|X_2 = 1)$ are $M_0 = \{\theta_{10}\theta_{210}, \theta_{11}\theta_{301}\}$ and $M_1 = \{\theta_{10}\theta_{210}, \theta_{11}\theta_{311}\}$, respectively. Then, using (4), we have:

$$\begin{aligned} p_0(\Theta) &= \hat{P}(X_3 = 0|X_2 = 1) = \\ &c_{01}m_{01} + c_{02}m_{02} = c_{01}\theta_{10}\theta_{210} + c_{02}\theta_{11}\theta_{301}. \end{aligned} \tag{8}$$

$$\begin{aligned} p_1(\Theta) &= \hat{P}(X_3 = 1|X_2 = 1) = \\ &c_{11}m_{11} + c_{12}m_{12} = c_{11}\theta_{10}\theta_{210} + c_{12}\theta_{11}\theta_{311}. \end{aligned} \tag{9}$$

- **Step 9:** For each possible state j of node X_i, calculate the coefficients c_{jr} of the conditional probabilities in (4), $r = 0, \ldots, n_j$, as follows:
  1. Calculate n_j different instantiations of Θ, $C = \{\theta_1, \ldots, \theta_{n_j}\}$ such that the canonical $n_j \times n_j$ matrix $\mathbf{T}_j$, whose rs-th element is the value of the monomial m_r obtained by replacing Θ by θ_s, is a non-singular matrix.
  2. Use any numeric network inference method to compute the vector of numerical probabilities $\mathbf{p}_j = (p_j(\theta_1), \ldots, p_j(\theta_{n_j}))$ by propagating the evidence $E = e$ in the reduced graph D' obtained in Step 2.
  3. Calculate the vector of coefficients $\mathbf{c}_j = (c_{j1}, \ldots, c_{jn_j})$ by solving the system of equations

$$\mathbf{T}_j \mathbf{c}_j = \mathbf{p}_j. \tag{10}$$

- **Step 9 (example):** Thus, taking appropriate combinations of extreme values for the symbolic parameters (canonical components), we can obtain the numeric coefficients by propagating the evidence not in the original graph D (Castillo, Gutiérrez and Hadi 1996), but in the reduced graph D', saving a lot of computation time. We have the symbolic parameters $\Theta = (\theta_{10}, \theta_{11}, \theta_{200}, \theta_{210}, \theta_{301}, \theta_{311})$ contained in D', We take the canonical components $\theta_1 = (1, 0, 1, 0, 1, 0)$ and $\theta_2 = (0, 1, 0, 1, 1, 0)$ and using any (exact or approximate) numeric network inference methods to calculate the coefficients of $p_0(\Theta)$. We obtain, $p_0(\theta_1) = 0.4$ and $p_0(\theta_2) = 0.3$. Note that, in the above equation, the vector $(p_0(\theta_1), p_0(\theta_2))$ can be calculated using any of the standard exact or approximate numeric network inference methods, because all the symbolic parameters have been assigned a numerical value:

$$\begin{aligned} &(p_0(\theta_1), p_0(\theta_2)) = (P(X_3 = 0|X_2 = 1, \Theta = \theta_1), \\ &P(X_3 = 0|X_2 = 1, \Theta = \theta_2)). \end{aligned}$$

Then, no symbolic operations are performed to obtain the symbolic solution. Thus, (10) becomes

$$\begin{pmatrix} c_{01} \\ c_{02} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p_0(\theta_1) \\ p_0(\theta_2) \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.3 \end{pmatrix}. \tag{11}$$

Similarly, taking the canonical components $\theta_1 = (1, 0, 1, 0, 1, 0)$ and $\theta_2 = (0, 1, 0, 1, 0, 1)$, for the probability $p_1(\Theta)$ we obtain

$$\begin{pmatrix} c_{11} \\ c_{12} \end{pmatrix} = \begin{pmatrix} 0.6 \\ 0.3 \end{pmatrix}. \tag{12}$$

Then, by substituting in (8) and (9), we obtain the unnormalized probabilities:

$$\hat{P}(X_3 = 0|X_2 = 1) = 0.4\theta_{10}\theta_{210} + 0.3\theta_{11}\theta_{301}. \quad (13)$$

$$\hat{P}(X_3 = 1|X_2 = 1) = 0.6\theta_{10}\theta_{210} + 0.3\theta_{11}\theta_{311}. \quad (14)$$

- **Step 10:** Calculate the unnormalized probabilities $p_j(\Theta)$, $j = 0, \ldots, r_i$ and the conditional probabilities $P(X_i = j|E = e) = p_j(\Theta)/N$, where

$$N = \sum_{j=0}^{r_i} p_j(\Theta)$$

is the normalizing constant.

- **Step 10 (example):** Finally, normalizing (13) and (14) we get the final polynomial expressions:

$$P(X_3 = 0|X_2 = 1) = \frac{0.4\theta_{10}\theta_{210} + 0.3\theta_{11}\theta_{301}}{\theta_{10}\theta_{210} + 0.3\theta_{11}\theta_{301} + 0.3\theta_{11}\theta_{311}} \quad (15)$$

and

$$P(X_3 = 1|X_1 = 1) = \frac{0.6\theta_{10}\theta_{210} + 0.3\theta_{11}\theta_{311}}{\theta_{10}\theta_{210} + 0.3\theta_{11}\theta_{301} + 0.3\theta_{11}\theta_{311}}. \quad (16)$$

- **Step 11:** Use (3) to eliminate dependent parameters and obtain the final expression for the conditional probabilities.
- **Step 11 (example):** Now, we apply the relationships among the parameters in (3) to simplify the above expressions. In this case, we have: $\theta_{311} = 1-\theta_{301}$ and $\theta_{11} = 1-\theta_{10}$. Thus, we get Expressions (5) and (6).
  Equations (5) and (6) give the posterior distribution of the goal node X_3 given the evidence $X_2 = 1$ in symbolic form. Thus, $P(X_3 = j|X_2 = 1), j = 0, 1$ can be evaluated directly by plugging in (5) and (6) any specific combination of values for the symbolic parameters without the need to redo the propagation from scratch for every given combination of values.

Remark: In some cases, it is possible to obtain a set of canonical instantiations for the above algorithm that leads to an identity matrix $\mathbf{T}_j$. In those cases, the coefficients of the symbolic expressions are directly obtained from numeric network inferences, without the extra effort of solving a system of linear equations. This fact is illustrated in the following example.

Sensitivity Analysis

The lower and upper bound of the resulting symbolic expressions are a useful information for performing sensitivity analysis (Castillo, Gutiérrez and Hadi 1996a). In this section we show how to obtain an interval, $(l, u) \subset [0, 1]$, that contains all the solutions of the given problem, for any combination of numerical values for the symbolic parameters. These bounds of the obtained ratios of polynomials as, for example (5) and (6), are attained at one of the canonical components (vertices of the feasible convex parameter set). We use the following theorem given by Martos 1964.

θ_k			$P(X_3 = j\|X_2 = 1, \theta_k)$	
θ_{10}	θ_{210}	θ_{301}	$j = 0$	$j = 1$
0	0	0	0.0	1.0
0	0	1	1.0	0.0
0	1	0	0.0	1.0
0	1	1	1.0	0.0
1	0	0	0.4	0.6
1	0	1	0.4	0.6
1	1	0	0.4	0.6
1	1	1	0.4	0.6

Table 2: Conditional probabilities for the canonical cases associated with $\theta_{10}, \theta_{210}$, and θ_{301}.

Theorem 1 If the linear fractional functional of a vector u,

$$\frac{\mathbf{c} * \mathbf{u} - c_0}{\mathbf{d} * \mathbf{u} - d_0}, \quad (17)$$

where $\mathbf{c}$ and $\mathbf{d}$ are vector coefficients and c_0 and d_0 are real constants, is defined in the convex polyhedron $A\mathbf{u} \leq a_0, \mathbf{u} \geq 0$, where A is a constant matrix and a_0 is a constant vector, and the denominator in (17) does not vanish in the polyhedron, then the functional reaches the maximum at least in one of the vertices of the polyhedron. ■

In our case, $\mathbf{u}$ is the set of symbolic parameters and the fractional functions (17) are the symbolic expressions associated with the probabilities, (5) and (6). In this case, the convex polyhedron is defined by $\mathbf{u} \leq 1, \mathbf{u} \geq 0$, that is, $\mathbf{A}$ is the identity matrix. Then, using Theorem 1, we know that the lower and upper bounds of the symbolic expressions associated with the probabilities are attained at the vertices of this polyhedron. In our case, the vertices of the polyhedron are given by all possible combinations of values 0 or 1 of the symbolic parameters, that is, by the complete set of canonical components associated with the set of free symbolic parameters appearing in the final symbolic expressions.

As an example, Table 2 shows the canonical probabilities associated with the symbolic expressions (5) and (6) obtained for the conditional probabilities $P(X_3 = j|X_2 = 1)$. The minimum and maximum of these probabilities is 0 and 1, respectively. Therefore, the lower and upper bounds are trivial bounds in this case. The same trivial bounds are obtained when fixing the symbolic parameters θ_{10} or θ_{210} to a given numeric parameter.

However, if we consider a numeric value for the symbolic parameter θ_{301}, for example $\theta_{301} = 0.5$, we obtain the canonical probabilities shown in Table 3. Therefore, the lower and upper bounds for the probability $P(X_3 = 0|X_2 = 1)$ in this situation are $(0.4, 0.5)$, and for $P(X_3 = 1|X_2 = 1)$ are $(0.5, 0.6)$, both with a range of 0.1.

θ_k		$P(X_3 = j \mid X_2 = 1, \theta_k)$	
θ_{10}	θ_{210}	$j = 0$	$j = 1$
0	0	0.5	0.5
0	1	0.5	0.5
1	0	0.4	0.6
1	1	0.4	0.6

Table 3: Conditional probabilities for the canonical cases associated with θ_{10} and θ_{210} for $\theta_{301} = 0.5$.

If we instantiate another symbolic parameter, for example $\theta_{10} = 0.1$, the new range is lower than in the previous case. We obtain the lower and upper bounds $(0.473, 0.5)$ for $P(X_3 = 0|X_2 = 1)$, and $(0.5, 0.537)$ for $P(X_3 = 1|X_2 = 1)$.

Conclusions and Recommendations

We have presented an efficient goal oriented algorithm for symbolic propagation in Bayesian networks, which allows dealing with symbolic or mixed cases of symbolic-numeric parameters. The main advantage of this algorithm is that uses numeric network inference methods, which allow it to compete with an impressive advantage with pure symbolic methods. First, the initial graph is reduced to produce a new graph which contains only the relevant nodes and parameters required to compute the desired propagation. Next, the relevant monomials in the symbolic parameters appearing in the target probabilities are identified. Then, the symbolic expression of the solution is obtained by performing numerical propagations associated with specific numerical values (canonical components) of the symbolic parameters. Furthermore, the canonical components allow us to obtain lower and upper bounds for the symbolic marginal or conditional probabilities. An example is used to illustrate the proposed methodology.

Acknowledgments

We thank the Dirección General de Investigación Científica y Técnica (DGICYT) (project PB94-1056) and Iberdrola for partial support of this work.

References

Bouckaert, R. R., Castillo, E. and Gutiérrez, J. M. 1995. A Modified Simulation Scheme for Inference in Bayesian Networks. *International Journal of Approximate Reasoning*. To appear.

Castillo, E., Gutiérrez, J. M., and Hadi, A. S. 1995. Parametric Structure of Probabilities in Bayesian Networks. *Lectures Notes in Artificial Intelligence*, Springer-Verlag, 946:89-98.

Castillo, E., Gutiérrez, J. M., and Hadi, A. S. 1996a. A New Method for Efficient Symbolic Propagation in Discrete Bayesian Networks. *Networks*. To appear.

Castillo, E., Gutiérrez, J. M., and Hadi, A. S. 1996b. *Expert Systems and Probabilistic Network Models*. Springer-Verlag, New York.

Fung, R. and Chang, K-C. 1990. Weighing and Integrating Evidence for Stochastic Simulation in Bayesian Networks, in *Uncertainty in Artificial Intelligence 5*, Machine Intelligence and Pattern Recognition Series, 10, (Henrion et al. Eds.), North Holland, Amsterdam, 209-219.

Geiger, D., Verma, T., and Pearl, J. 1990. Identifying Independence in Bayesian Networks. *Networks*, 20:507–534.

Henrion, M. 1988. Propagating Uncertainty in Bayesian Networks by Probabilistic Logic Sampling, in *Uncertainty in Artificial Intelligence*, (J.F. Lemmer and L. N. Kanal, Eds.), North Holland, Amsterdam, 2:317-324.

Lauritzen, S. L. and Spiegelhalter, D. J. 1988. Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. *Journal of the Royal Statistical Society (B)*, 50:157-224.

Li, Z., and D'Ambrosio, B. 1994. Efficient Inference in Bayes Nets as a Combinatorial Optimization Problem. *International Journal of Approximate Reasoning*, 11(1):55-81.

Martos, B. 1964. Hyperbolic Programming. *Naval Research Logistic Quarterly*, 32:135-156.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA.

Poole, D. 1993. Average-case Analysis of a Search Algorithm for Estimating Prior and Posterior Probabilities in Bayesian Networks with Extreme Probabilities, in *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, 13(1):606-612.

Shachter, R. D. 1990. An Ordered Examination of Influence Diagrams. *Networks*, 20:535-563.

Shachter, R. D., D'Ambrosio, B., and DelFabero, B. 1990. Symbolic Probabilistic Inference in Belief Networks, in *Proceedings Eighth National Conference on AI*, 126-131.

Shachter, R. D. and Peot, M. A. 1990. Simulation Approaches to General Probabilistic Inference on Belief Networks, in *Uncertainty in Artificial Intelligence*, Machine Intelligence and Pattern Recognition Series, 10 (Henrion et al. Eds.), North Holland, Amsterdam, 5:221-231.