

# Parametric Structure of Probabilities in Bayesian Networks

Enrique Castillo<sup>1</sup>, José Manuel Gutiérrez<sup>1</sup> and Ali S. Hadi<sup>2</sup>

<sup>1</sup> Department of Applied Mathematics and Computational Science,  
University of Cantabria, SPAIN

<sup>2</sup> Department of Statistics, Cornell University, USA

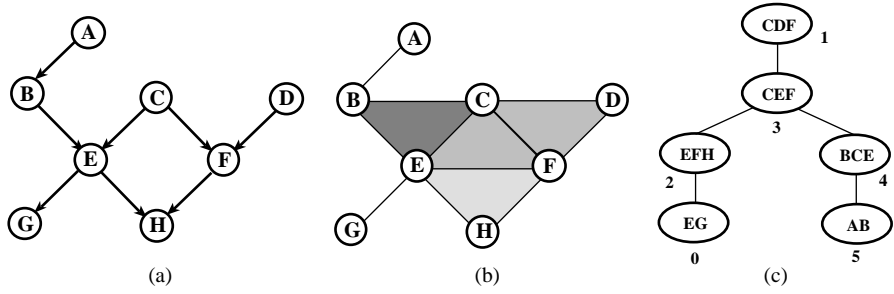
**Abstract.** The paper presents a method for uncertainty propagation in Bayesian networks in symbolic, as opposed to numeric, form. The algebraic structure of probabilities is characterized. The prior probabilities of instantiations and the marginal probabilities are shown to be rational functions of the parameters, where the polynomials appearing in the numerator and the denominator are at the most first degree in each of the parameters. It is shown that numeric propagation algorithms can be adapted for symbolic computations by means of canonical components. Furthermore, the same algorithms can be used to build automatic code generators for symbolic propagation of evidence. An example of uncertainty propagation in a clique tree is used to illustrate all the steps and the corresponding code in *Mathematica* is given. Finally, it is shown that upper and lower bounds for the marginal probabilities of nodes are attained at one of the canonical components.

## 1 Introduction

Bayesian networks are powerful tools for handling uncertainty in expert systems. A key problem in Bayesian networks is evidence propagation. There are several well-known methods for exact and approximate propagation of evidence in a Bayesian network; see, for example, Pearl [1, 2], Lauritzen and Spiegelhalter [3], Castillo and Alvarez [4], and Castillo, Gutiérrez and Hadi [5, 6]. These methods, however, require that the joint probabilities of the nodes be given in a numeric form. In practice, exact numeric specification of these parameters may not be available. In such cases, there is a need for methods which are able to deal with the parameters symbolically. Symbolic propagation leads to probabilities which are expressed as functions of the parameters instead of real numbers. Thus, the answers to specific queries can then be obtained by plugging the values of the parameters in the solution, without need to redo the propagation. Furthermore, a real practical use of this approach is the possibility of performing a sensitivity analysis of the parameter values without the need of redoing the computations.

## 2 Notation and Basic Framework

Let  $X = \{X_1, X_2, \dots, X_n\}$  be a set of  $n$  discrete variables and let  $r_i$  be the cardinality (number of states) of variable  $X_i$ . A Bayesian network  $B$  over  $X$  is a



**Fig. 1.** Example of a Bayesian Network with its cliques and associated clique tree.

pair  $(B_S, B_P)$ , where the network structure  $B_S$  is a directed acyclic graph with one node for each variable in  $X$ . We refer to the parents of a node  $X_i$  in the network structure as  $\Pi_i$  and to the set of instances of  $\Pi_i$  (specific combination of values in the parent set) as  $\pi_i$  (in lower case).  $B_P$  is a set of  $n$  conditional probabilities  $P_i(X_i|\Pi_i)$ , one for each variable, which gives the probabilities of  $X_i$ , given the values of the variables in its parent set  $\Pi_i$ . Using the chain rule, the joint probability density of  $X$  can be written as

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P_i(X_i|\Pi_i). \quad (1)$$

An advantage of this decomposition of the joint probability density is that each conditional probability  $P_i(X_i|\Pi_i)$  can be given independently of the other probabilities. We can consider  $P_i(X_i|\Pi_i)$  as a parametric family. A natural choice of parameters is given by

$$\theta_{ijk} = P_i(k|j), \quad k \in \{1, \dots, r_i-1\}, \quad 0 \leq j \leq q_i; \quad \sum_{k=1}^{r_i-1} \theta_{ijk} \leq 1, \quad i \in \{1, \dots, n\}, \quad (2)$$

where  $q_i = \prod_{i \in I_{\Pi_i}} r_i$  is the cardinal of  $\Pi_i$ ,  $I_{\Pi_i}$  is the set of indices of  $\Pi_i$ , and

$$\theta_{ijr_i} = 1 - \sum_{k=1}^{r_i-1} \theta_{ijk}. \quad (3)$$

For illustrative purpose, we shall use the following example:

*Example 1.* Assume a Bayesian network  $(B_S, B_P)$  where  $B_S$  is the directed graph in Fig. 1(a) with the corresponding cliques in Fig. 1(b) and associated clique tree in Fig. 1(c), which imply a joint probability of the set of nodes  $X$  of the form:

$$P(X) = P(A)P(B|A)P(C)P(D)P(E|BC)P(F|CD)P(G|E)P(H|EF). \quad (4)$$

Assume that all nodes represent binary variables with values in the set  $\{0, 1\}$  and that the conditional probabilities, numeric and symbolic, are given in Table 1.

$P(A = 0) = p_0$	$P(B = 0 A = 0) = 0.3$
$P(C = 0) = 0.4$	$P(B = 0 A = 1) = 0.5$
$P(E = 0 B = 0, C = 0) = 0.2$	$P(F = 0 C = 0, D = 0) = 0.1$
$P(E = 0 B = 0, C = 1) = 0.3$	$P(F = 0 C = 0, D = 1) = 0.4$
$P(E = 0 B = 1, C = 0) = 0.4$	$P(F = 0 C = 1, D = 0) = 0.3$
$P(E = 0 B = 1, C = 1) = 0.5$	$P(F = 0 C = 1, D = 1) = 0.2$
$P(D = 0) = p_1$	$P(H = 0 E = 0, F = 0) = 0.2$
$P(G = 0 E = 0) = 0.3$	$P(H = 0 E = 0, F = 1) = 0.4$
$P(G = 0 E = 1) = 0.6$	$P(H = 0 E = 1, F = 0) = 0.6$
	$P(H = 0 E = 1, F = 1) = 0.3$

**Table 1.** Conditional probability tables showing two symbolic parameters:  $p_0$  and  $p_1$ .

### 3 Exact Propagation in Bayesian Networks

In this section we consider an exact method for the propagation of uncertainties (see, for example, Shachter, Andersen, and Szolovits [7]). This method will serve as the basis for the symbolic methods in the following sections. Initially, each conditional probability distribution  $P(X_i|\Pi_i)$  is assigned to exactly one clique containing the  $X_i$ 's family,  $(X_i, \Pi_i)$ . The product of the conditional probability distributions assigned to clique  $S_i$  is called the *potential function*,  $\Psi_i(S_i)$ , for clique  $S_i$ . With this, once the evidence  $\mathcal{E}$  is known, the joint distribution for all the variables can be written as

$$P(X) = \prod_{X_i \in X - \mathcal{E}} P(X_i|\Pi_i) \prod_{e_k \in \mathcal{E}} P(e_k|\Pi_k) = \prod_{k=1}^{N_c} \Psi_k(S_k), \quad (5)$$

where  $N_c$  is the number of cliques and  $e_k$  is the evidence. Shachter, Andersen, and Szolovits [7] show that the joint probability for clique  $S_i$  is

$$P_i(S_i) = \Psi_i(S_i) \prod_{k \in A_i} M_{ki}, \quad (6)$$

where  $A_i$  is the set of all cliques adjacent to  $S_i$  and  $M_{ki}$  is the message sent by clique  $S_k$  to  $S_i$ , which is given by:

$$M_{ij} = \left( \sum_{(S_i - S_j)} \Psi_i(S_i) \prod_{k \in A_i - j} M_{ki} \right). \quad (7)$$

Once  $P_i(S_i)$  is known, the node marginals can be easily calculated by marginalizing in the cliques. These expressions lead to well-known numerical propagation algorithms.

## 4 Symbolic Computations

Dealing with symbolic computations is the same as dealing with numeric values with the only difference being that all the required operations must be performed by a program with symbolic manipulation capabilities. Symbolic computations, however, are intrinsically slow and require more memory. A code generator can be easily written based on any standard propagation algorithm if, instead of building the potential functions  $\Psi_i(S_i)$ , and calculating the  $M_{ij}$  messages, the probability function of the cliques, and the node marginals, we write the corresponding code in the order indicated by the algorithm. We have written such a program in C++ language. In fact the code in Fig. 2 has been generated by this computer program given the network in Example 1. Table 2 shows that the initial marginal probabilities of the nodes are polynomials in the parameters.

Node	P(Node=0)	Node	P(Node=0)
A	$p_0$	B	$0.5 - 0.2p_0$
C	0.4	D	$p_1$
E	$0.04(9 + p_0)$	F	$0.02(14 - 3p_1)$
G	$0.012(41 - p_0)$	H	$0.0004(930 - 4p_0 - 30p_1 + 3p_0p_1)$

**Table 2.** Initial probabilities of nodes.

The previous results have been obtained without any instantiation of evidence. Suppose now we have the following evidence:  $\{D = 0, E = 0\}$ . To this aim, we make the ranges of variables  $D$  and  $E$  equal to  $(0, 0)$ , that is,  $u1[3] = u1[4] = 0$  and repeat the calculations. Table 3 gives the new probabilities of the nodes given this evidence. We get rational functions, i.e., quotients of polynomial functions in the parameters with unit exponents. The fact that the probability of any instantiation is a polynomial in the parameters is proven in Sect. 5.

Node	P(Node=0)	Node	P(Node=0)
A	$(10p_0)/(9 + p_0)$	B	$0.65(5 - 2p_0)/(9 + p_0)$
C	$0.2(15 + 2p_0)/(9 + p_0)$	D	1
E	1	F	$0.02(105 + 11p_0)/(9 + p_0)$
G	0.3	H	$0.004(795 + 89p_0)/(9 + p_0)$

**Table 3.** Conditional probabilities of the nodes, given  $\{D = 0, E = 0\}$ , showing the common denominator

Tables 2 and 3 can then be used to answer all queries regarding initial or ev-

```

(* Probability Tables and Initialize ranges*)
T={p0,1-p0};n=1; Do[PA[i1]=T[[n]];n++,{i1,0,1}];
T={0.3,0.5,0.7,0.5};n=1; Do[PB[i1,i2]=T[[n]];n++,{i1,0,1},{i2,0,1}];
T={0.4,0.6};n=1; Do[PC[i1]=T[[n]];n++,{i1,0,1}];
T={p1,1-p1};n=1; Do[PD[i1]=T[[n]];n++,{i1,0,1}];
T={0.2,0.3,0.4,0.5,0.8,0.7,0.6,0.5};
n=1; Do[PE[i1,i2,i3]=T[[n]];n++,{i1,0,1},{i2,0,1},{i3,0,1}];
T={0.1,0.4,0.3,0.2,0.9,0.6,0.7,0.8};
n=1;Do[PF[i1,i2,i3]=T[[n]];n++,{i1,0,1},{i2,0,1},{i3,0,1}];
T={0.3,0.6,0.7,0.4}; n=1; Do[PG[i1,i2]=T[[n]];n++,{i1,0,1},{i2,0,1}];
T={0.2,0.4,0.6,0.3,0.8,0.6,0.4,0.7};
n=1; Do[PH[i1,i2,i3]=T[[n]];n++,{i1,0,1},{i2,0,1},{i3,0,1}];
Do[u0[i]=0; u1[i]=1, {i,0,7}];

(* Potential Functions *)
F0[G_,E_]:=PG[G,E];F1[D_,C_,F_]:=PC[C]*PD[D]*PF[F,C,D];
F2[H_,E_,F_]:=PH[H,E,F];F3[F_,E_,C_]:=1;
F4[C_,B_,E_]:=PE[E,B,C];F5[B_,A_]:=PA[A]*PB[B,A];

(* Messages *)
L02[E_]:=Sum[F0[G,E],{G,u0[6],u1[6]}];
L13[C_,F_]:=Sum[F1[D,C,F],{D,u0[3],u1[3]}];
L23[E_,F_]:=Sum[F2[H,E,F]*L02[E],{H,u0[7],u1[7]}];
L34[E_,C_]:=Sum[F3[F,E,C]*L13[C,F]*L23[E,F],{F,u0[5],u1[5]}];
L45[B_]:=Sum[F4[C,B,E]*L34[E,C],{C,u0[2],u1[2]},{E,u0[4],u1[4]}];
L54[B_]:=Sum[F5[B,A],{A,u0[0],u1[0]}];
L43[E_,C_]:=Sum[F4[C,B,E]*L54[B],{B,u0[1],u1[1]}];
L31[C_,F_]:=Sum[F3[F,E,C]*L23[E,F]*L43[E,C],{E,u0[4],u1[4]}];
L32[E_,F_]:=Sum[F3[F,E,C]*L13[C,F]*L43[E,C],{C,u0[2],u1[2]}];
L20[E_]:=Sum[F2[H,E,F]*L32[E,F],{H,u0[7],u1[7]},{F,u0[5],u1[5]}];

(* Cluster and Node Marginals *)
Q0[G_,E_]:=F0[G,E]*L20[E];Q1[D_,C_,F_]:=F1[D,C,F]*L31[C,F];
Q2[H_,E_,F_]:=F2[H,E,F]*L02[E]*L32[E,F];
Q3[F_,E_,C_]:=F3[F,E,C]*L13[C,F]*L23[E,F]*L43[E,C];
Q4[C_,B_,E_]:=F4[C,B,E]*L34[E,C]*L54[B];Q5[B_,A_]:=F5[B,A]*L45[B];
M[0,A_]:=Sum[Q5[B,A],{B,u0[1],u1[1]}];M[1,B_]:=Sum[Q5[B,A],{A,u0[0],u1[0]}];
M[2,C_]:=Sum[Q4[C,B,E],{B,u0[1],u1[1]},{E,u0[4],u1[4]}];
M[3,D_]:=Sum[Q1[D,C,F],{C,u0[2],u1[2]},{F,u0[5],u1[5]}];
M[4,E_]:=Sum[Q0[G,E],{G,u0[6],u1[6]}];M[6,G_]:=Sum[Q0[G,E],{E,u0[4],u1[4]}];
M[5,F_]:=Sum[Q3[F,E,C],{E,u0[4],u1[4]},{C,u0[2],u1[2]}];
M[7,H_]:=Sum[Q2[H,E,F],{E,u0[4],u1[4]},{F,u0[5],u1[5]}];

(* Normalizations *)
Do[sumConst=Chop[Simplify[Sum[M[i,t],{t,u0[i],u1[i]}]]];Do[R[i,t]:=
Simplify[Chop[M[i,t]/sumConst];Print["P(Node",i,"=",t,")="],R[i,t],
{t,u0[i],u1[i]}];Print[" "],{i,0,7}

```

**Fig. 2.** Mathematica statements for symbolic propagation of evidence.

identical marginal probabilities associated with the network in Example 1 simply by plugging in specific values for the parameters.

We note that the symbolic part of computations increases exponentially with the number of parameters but not with the number of nodes.

## 5 Algebraic Structure of Probabilities

In this section we discuss the algebraic structure of probabilities of single nodes. We start with the prior and later we analyze the case of posterior probabilities.

**Theorem 1.** *The prior probability of any instantiation is a polynomial in the parameters of degree less than or equal to the minimum of the number of parameters or nodes. However, it is a first degree polynomial in each parameter.*

*Proof.* According to (1) the probability of an instantiation  $(x_1, \dots, x_n)$  is

$$\prod_{i=1}^n P(x_i | \pi_i),$$

that is, a product of  $n$  factors. Each factor is either  $\theta_{ijk}$ , if  $x_i < r_i$  or

$$1 - \sum_{k=1}^{r_i-1} \theta_{ijk},$$

if  $x_i = r_i$  (see (3)), that is, a parameter or a first degree polynomial in some parameters.

In addition, each parameter appears at most in one factor and dependent parameters, such as  $\theta_{ijk_1}$  and  $\theta_{ijk_2}$ , do not appear in the same factor. Thus, we get a polynomial of degree less than or equal to the minimum of the number of parameters or nodes, which is first degree in each parameter.  $\square$

**Corollary 2.** *The prior node marginals are polynomials in the parameters of the same form.*

*Proof.* The prior marginals of any node are the sum of the probabilities of a subset of instantiations.  $\square$

It is well-known from probability theory, that after some evidence is available, the joint probability of the remaining nodes is proportional to (1) with the evidential nodes instantiated to their evidence values. Thus, while Equation (1) gives the true (normalized) joint probability of nodes, instantiation of the evidential variables in (1) leads to the unnormalized joint probability of the remaining nodes. Thus, the same methods can be used for prior and posterior probabilities, the only difference being that in the later case the normalization constant must be determined. However, this constant is common to the joint and to any of the possible marginal probabilities. Thus, we have the following corollary.

**Corollary 3.** *The posterior node marginals, i.e., the node marginals given some evidence  $\mathcal{E}$ , are rational functions of the parameters, that is, quotient of polynomials in the parameters of the same form. The denominator polynomial is the same for all nodes.*

*Proof.* When normalizing probabilities we divide by their sum, that is, by a polynomial of the same form; then, the rational functions arise.  $\square$

Because the denominator polynomial is the same for any of the possible marginals, for implementation purposes, it is more convenient to calculate and store all the numerator polynomials for each node and calculate and store the common denominator polynomial separately.

It is interesting to know the total number of monomials involved. This is given by the following theorem.

**Theorem 4.** *The total number of monomials is given by*

$$\prod_{i=1}^n (1 + s_i), \quad (8)$$

where  $s_i$  is the number of parameters  $\theta_{ijk}$ , i.e., those related to node  $i$ .

*Proof.* Each monomial is generated from  $n$  factors, each associated with a given node. Each factor can be a constant value or one of the parameters.  $\square$

**Corollary 5.** *The maximum number of monomials is given by*

$$\prod_{i=1}^n [1 + q_i(r_i - 1)]. \quad (9)$$

*Proof.* The maximum number of parameters related to node  $i$  is  $q_i(r_i - 1)$  and using (8), (9) holds.  $\square$

Once we know the structure of the marginal probabilities we can exploit it to obtain symbolic results using numerical procedures, as follows.

## 6 Symbolic Propagation and Numeric Methods

In this section we show how symbolic propagation can be performed using numeric methods. This has special importance for large networks. Consider the network in Example 1 and assume that we want to know the influence of the parameters  $p_0$  and  $p_1$  on the conditional probabilities of the remaining nodes given the evidence  $\{D = 0, E = 0\}$ . Then, for any node  $i$  we know that the unnormalized marginal probability  $p_{ik}^*(p_0, p_1)$  is a polynomial of the form

$$P(X_i = k | D = 0, E = 0) = a_{ik} + b_{ik}p_0 + c_{ik}p_1 + d_{ik}p_0p_1 = p_{ik}^*(p_0, p_1). \quad (10)$$

Choosing  $C = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$ , which is the so-called *canonical component* set, we get the system of equations

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a_{ik} \\ c_{ik} \\ b_{ik} \\ d_{ik} \end{pmatrix} = \begin{pmatrix} p_{ik}^*(0, 0) \\ p_{ik}^*(0, 1) \\ p_{ik}^*(1, 0) \\ p_{ik}^*(1, 1) \end{pmatrix}, \quad (11)$$

from which the polynomial coefficients can be calculated.

In fact, we can use any set of  $(p_0, p_1)$ , normalized or unnormalized, values such that the leading matrix in (11) becomes non-singular. We use the fact that our probability can be written as a linear convex combination of a given generating set of probabilities (the canonical probabilities), i.e., it belongs to the convex hull generated by them.

Adding the unnormalized probabilities of any node we get the normalization polynomial (common denominator). Note that with this method we can deal with symbolic propagation using numeric programs. Note also that exact and approximate numeric algorithms can be used.

### 6.1 Computing Canonical Components

When propagating uncertainty in the canonical cases we can save many calculations because some messages are common to all of them. Fig. 3 shows all the clique messages indicated by arrows. We can distinguish three types of messages:

1. Messages with no index. These are common messages that need to be calculated only once.
2. Messages with only one index  $D$  or  $E$ . These messages depend on the parameters associated with the index node and then we must calculate as many different messages as the number of monomials associated with it.
3. Messages with two indices  $D$  and  $E$ . These messages depend on the parameters associated with nodes  $D$  and  $E$  and then we must calculate as many different messages as the number of monomials associated with them.

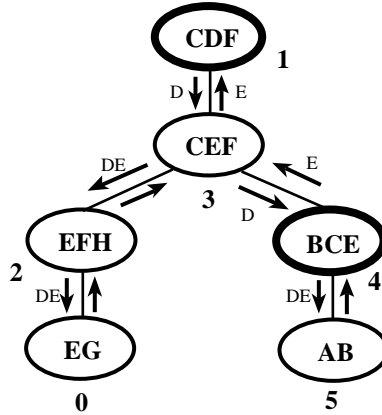
We can build the rational function associated with the node marginals based only on numeric calculations. Thus, we have a numeric method that efficiently solves the symbolic problem.

## 7 Symbolic Treatment of Random Evidence

Until now we have dealt with deterministic evidence. In this section we deal with random evidence. Let  $Q_{e_1, \dots, e_k}(x_i)$  denote the marginal of node  $X_i$  when the evidence nodes take values  $e_1, \dots, e_k$ , where  $k$  is the cardinality of  $\mathcal{E}$ . If we assume a lottery  $q(e_1, \dots, e_k)$ , that is, a probability over  $\mathcal{E}$ , then the marginal of node  $X_i$  becomes

$$P_q(x_i) = \sum_{e_1, \dots, e_k} q(e_1, \dots, e_k) Q_{e_1, \dots, e_k}(x_i); \quad \sum_{e_1, \dots, e_k} q(e_1, \dots, e_k) = 1. \quad (12)$$





**Fig. 3.** Clique tree and messages affected by the parameters in nodes  $D$  and  $E$ . Cliques involving potential functions depending on parameters  $p_0$  and  $p_1$  are outlined.

It is important to note that  $P_q(x_i)$  is also a rational function because it is a linear convex combination of the rational functions  $Q_{e_1, \dots, e_k}(x_i)$ . However, in this case the conditional probability table parameters can appear with exponents larger than one, which implies polynomial of order larger than one in each of the parameters. The following theorem states this result.

**Theorem 6.** *The probabilities of the nodes given a random evidence are rational functions where the degree in each parameter of the polynomials involved is at most the sum of the cardinalities of the evidential nodes.*

*Proof.* The polynomial denominators of the rational functions  $Q_{e_1, \dots, e_k}(x_i)$  are in general different for different combinations of the evidence set  $(e_1, \dots, e_k)$ . Thus, reduction to common denominator to get the rational function (12) implies their product and then the result.  $\square$

As an example, the probabilities of node  $C = 0$  for the random evidence  $P(F = 0) = a$  is given by

$$a P(C = 0|F = 0) + (1-a) P(C = 0|F = 1) = \frac{2(3p_1^2 + 40ap_1 - 8p_1 - 20a - 28)}{3p_1^2 + 22p_1 - 168}.$$

## 8 Upper and Lower Bounds for Probabilities

Symbolic expressions, such as those in Tables 2 and 3, can also be used to obtain upper and lower bounds for the marginal probabilities which is a valuable information. This can be easily done by considering non-informative  $\{0, 1\}$  bounds or other bounds given by experts. Using a theorem given by Bela-Martos (see [8]) it is immediate to see that upper and lower bounds are attained at one of

the canonical components (vertices of the feasible convex parameter set). As an example, from Table 2 the maximum and minimum values for the initial probabilities of node  $H = 0$  are attained in the set  $\{0.372, 0.36, 0.3704, 0.3596\}$ .

## 9 Conclusions

The symbolic structure of prior and posterior marginal probabilities of Bayesian networks have been characterized as polynomials or rational functions of the parameters, respectively, and the degrees of the polynomials have been shown to be dependent on the deterministic or random character of evidence. This characterization allows the symbolic propagation of evidence to be converted to a numeric problem using the canonical components, leading to an important saving in computation. An extra saving is obtained by identifying those messages which are common for all or some canonical components. These components are also shown to attain upper and lower bounds for the probabilities.

## Acknowledgements

We thank the University of Cantabria and Dirección General de Investigación Científica y Técnica (DGICYT) (project PB92-0504), for support of this work.

## References

1. Pearl, J.: Fusion, Propagation and Structuring in Belief Networks. *Artificial Intelligence*, **29** (1986) 241–288
2. Pearl, J.: Evidential Reasoning Using Stochastic Simulation of Causal Models. *Artificial Intelligence*, **32** (9186) 245–287
3. Lauritzen, S. L., Spiegelhalter, D. J.: Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. *Journal of the Royal Statistical Society (B)*, **50** (1988) 157–224
4. Castillo, E., Alvarez, E.: *Experts Systems: Uncertainty and Learning*. Computational Mechanics Publications and Elsevier Science. London (1991)
5. Castillo, E., Gutiérrez, J. M., Hadi, A. S.: *Modelling Probabilistic Networks Based on Conditional Probabilities*. Technical report No. 93-2, (1993) University of Cantabria.
6. Castillo, E., Gutiérrez, J. M., Hadi, A. S.: *Expert Systems and Bayesian Networks*, to appear.
7. Shachter, R. D., Andersen, S. K., Szolovits, P.: *Global Conditioning for Probabilistic Inference in Belief Networks*. Proceedings UAI, Morgan Kaufmann Publishers (1994) 514–522.
8. Martos, B.: Hyperbolic programming. *Naval Research Logistic Quarterly*, **11** (1964) 135–156