Cantabria University

EVIDENCE PROPAGATION

Propagation of evidence consists of updating the probability distributions of the variables according to the newly available evidence. For example, we need to calculate the conditional distribution of each element of a set of variables of interest (e.g., diseases) given the evidence (e.g., symptoms).

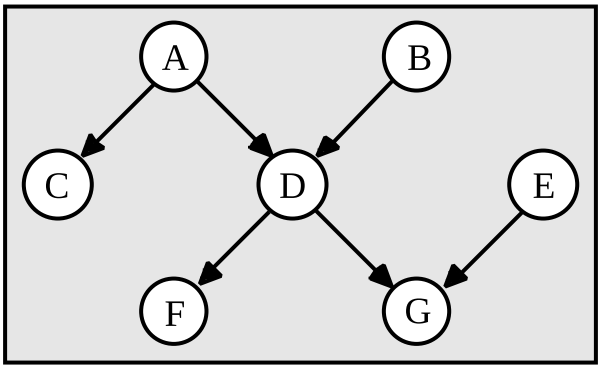
There are three types of algorithms for propagating evidence: exact, approximate, and symbolic.

Suppose we have a set of discrete variables $X = \{X_1, \ldots, X_n\}$ and a JPD, $p(x)$, over X. Before any evidence is available, the propagation process consists of calculating the marginal probability distribution (MPD) $p(X_i = x_i)$, or simply $p(x_i)$, for each $X_i \in X$.

Now, suppose that some evidence has become available, that is, a set of variables $E \subset X$ are known to take the values $X_i = e_i$, for $X_i \in E$ (**evidential variables**). In this situation, propagation of evidence consists of calculating the conditional probabilities $p(x_i|e)$.

Consider the graph in the figure below which implies

$$p(x) = p(a)p(b)p(c|a)p(d|a,b)p(e)p(f|d)p(g|d,e),$$



Brute-force method

$$p(d) = \sum_{x \setminus d} p(x) = \sum_{a,b,c,e,f,g} p(a,b,c,d,e,f,g).$$

Optimizing the summations

$$p(d) = \sum_{a,b,c,e,f,g} p(a)p(b)p(c|a)p(d|a,b)p(e)p(f|d)p(g|d,e)$$

$$= \left(\sum_{a,b,c} p(a)p(b)p(c|a)p(d|a,b)\right)\left(\sum_{e,f,g} p(e)p(g|d,e)p(f|d)\right),$$

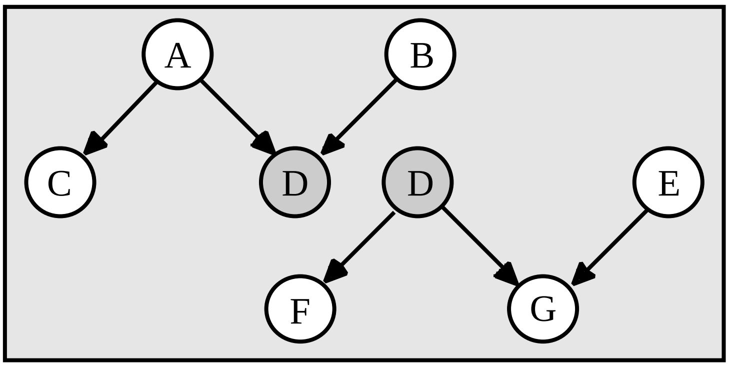
$$\sum_a \left[p(a)\sum_c \left[p(c|a)\sum_b p(b)p(d|a,b)\right]\right]\sum_e \left[p(e)\sum_f \left[p(f|d)\sum_g p(g|d,e)\right]\right].$$

Cantabria University

PROPAGATION IN POLYTREES

The evidence E can be decomposed into two subsets:

- E_i^+, the subset of E that can be accessed from X_i through its parents.
- E_i^-, the subset of E that can be accessed from X_i through its children.



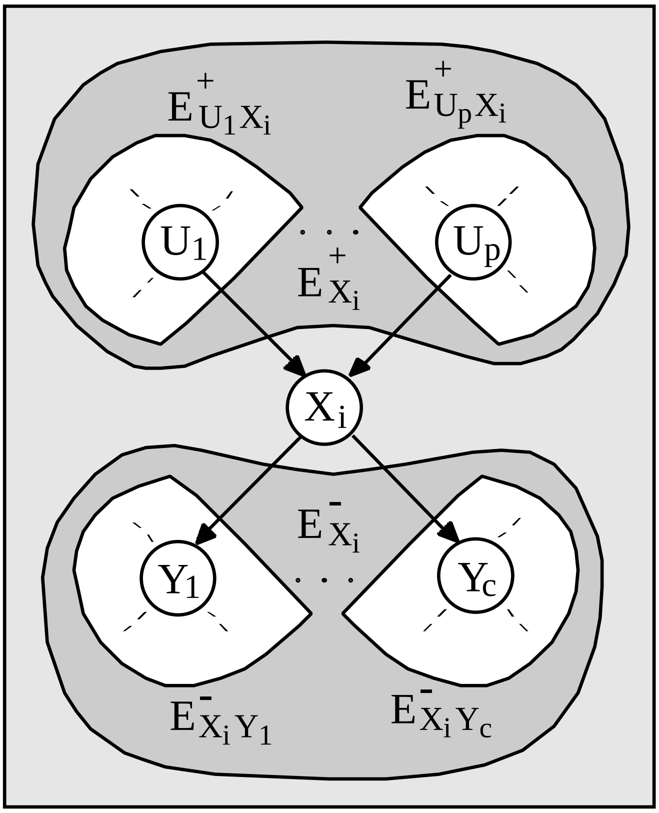
$$p(x_i|e) = p(x_i|e_i^-, e_i^+) = \frac{1}{p(e_i^-, e_i^+)} p(e_i^-, e_i^+|x_i)p(x_i).$$

Since X_i separates E_i^- from E_i^+ in the polytree, then the CIS $I(E_i^-, E_i^+|X_i)$ holds; hence we have

$$\begin{aligned} p(x_i|e) &= \frac{1}{p(e_i^-, e_i^+)} p(e_i^-|x_i)\, p(e_i^+|x_i)p(x_i) \\ &= \frac{1}{p(e_i^-, e_i^+)} p(e_i^-|x_i)\, p(x_i, e_i^+) \\ &= k\ \lambda_i(x_i)\ \rho_i(x_i), \end{aligned}$$

PROPAGATION IN POLYTREES

To compute the functions $\lambda_i(x_i)$ and $\rho_i(x_i)$, suppose that a typical node X_i has p parents, $U = \{U_1, \ldots, U_p\}$, and c children, $Y = \{Y_1, \ldots, Y_c\}$.



The evidence E_i^+ can be partitioned into p disjoint components, one for each parent of X_i:

$$E_i^+ = \{E^+_{U_1 X_i}, \ldots, E^+_{U_p X_i}\}, \qquad (1)$$

where the evidence $E^+_{U_j X_i}$ is the subset of E_i^+ contained in the U_j-side of the link $U_j \rightarrow X_i$.

Similarly, the evidence E_i^- can be partitioned into c disjoint components.

Cantabria University

SENDING MESSAGES

Let $u = \{u_1, \ldots, u_p\}$ be an instantiation of the parents of X_i:

$$\begin{aligned}\rho_i(x_i) &= p(x_i, e_i^+) = \sum_u p(x_i, u \cup e_i^+) \\ &= \sum_u p(x_i|u \cup e_i^+) p(u \cup e_i^+) \\ &= \sum_u p(x_i|u \cup e_i^+) p(u \cup e_{U_1 X_i}^+ \cup \ldots \cup e_{U_p X_i}^+).\end{aligned}$$

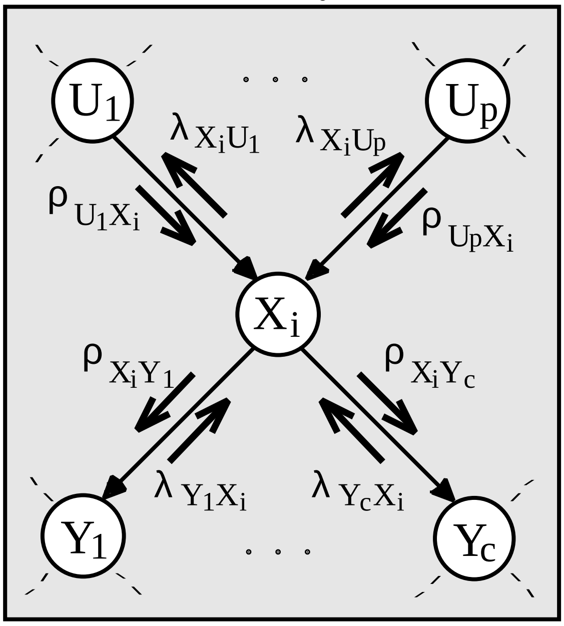
Due to the fact that $\{U_j, E_{U_j X_i}^+\}$ is independent of $\{U_k, E_{U_k X_i}^+\}$, for $j \neq k$, we have

$$\rho_i(x_i) = \sum_u p(x_i|u \cup e_i^+) \prod_{j=1}^{p} p(u_j \cup e_{U_j X_i}^+), \quad (2)$$

and

$$\rho_{U_j X_i}(u_j) = p(u_j \cup e_{U_j X_i}^+) \quad (3)$$

is the ρ-message that node U_j sends to its child X_i.



Cantabria University

PROPAGATION IN POLYTREES ALGORITHM

- **Input:** A Bayesian network model (D, P) over a set of variables X and a set of evidential nodes E with evidential values $E = e$, where D is a polytree.
- **Output:** The CPD $p(x_i|e)$ for every nonevidential node X_i.

EQUATIONS

$$\beta_i(x_i) = \lambda_i(x_i)\rho_i(x_i). \tag{4}$$

$$\rho_i(x_i) \sum_{u} p(x_i|u \cup e_i^+) \prod_{j=1}^{p} \rho_{U_j X_i}(u_j), \tag{5}$$

$$\lambda_i(x_i) = \prod_{j=1}^{c} \lambda_{Y_j X_i}(x_i), \tag{6}$$

where

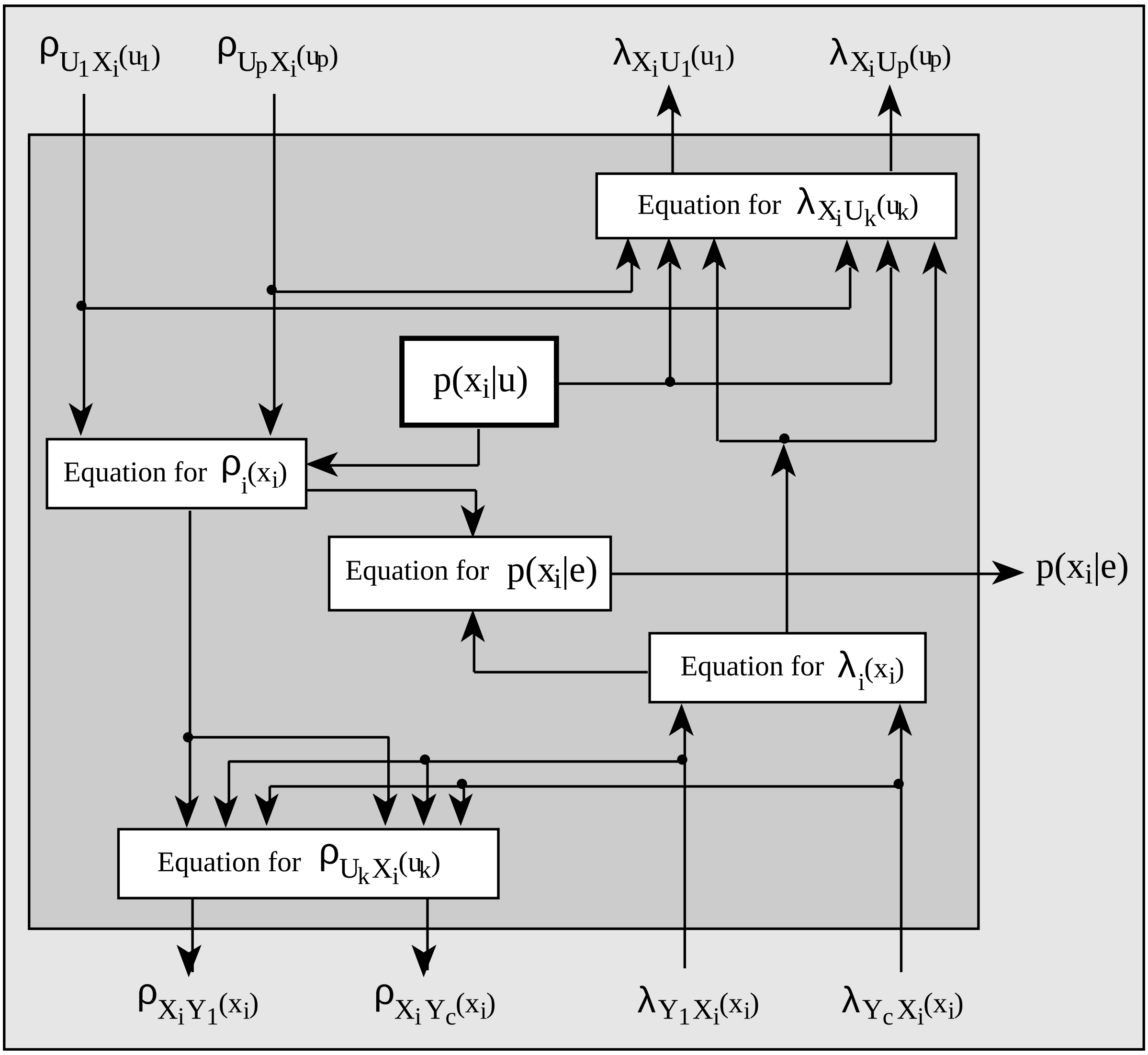
$$\lambda_{Y_j X_i}(x_i) = p(e^-_{X_i Y_j}|x_i) \tag{7}$$

$$\rho_{X_i Y_j}(x_i) \propto \rho_i(x_i) \prod_{k \neq j} \lambda_{Y_k X_i}(x_i). \tag{8}$$

$$\lambda_{Y_j X_i}(x_i) = \sum_{y_j} \lambda_{Y_j}(y_j) \sum_{v_1,\ldots,v_q} p(y_j|\pi_{Y_i}) \prod_{k=1}^{q} \rho_{V_k Y_j}(v_k). \tag{9}$$

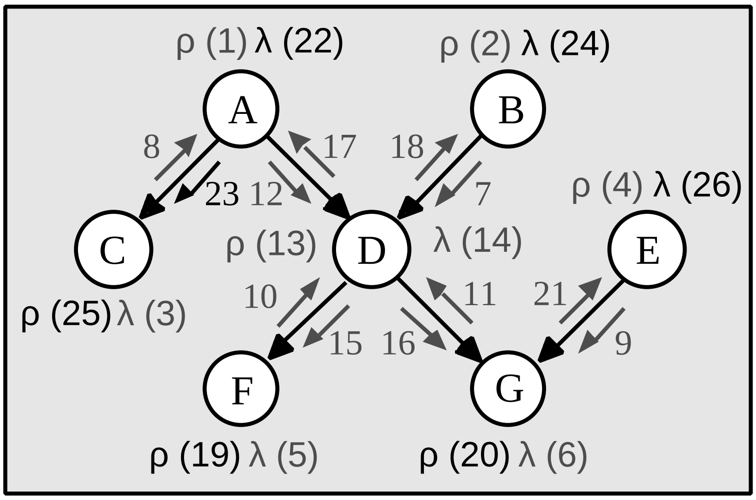
PARALLEL IMPLEMENTATION

The structure of message-passing makes the algorithm suitable for a parallel implementation. Assume that each node has its own processor.



Cantabria
University

AN EXAMPLE OF THE POLYTREES ALGORITHM



A				
a	ρ	λ	β	p(a)
0	0.3	1.0	0.3	0.3
1	0.7	1.0	0.7	0.7

B				
b	ρ	λ	β	p(b)
0	0.6	1.0	0.6	0.6
1	0.4	1.0	0.4	0.4

a	ρ_{AC}	λ_{CA}
0	0.3	1.0
1	0.7	1.0

a	ρ_{AD}	λ_{DA}
0	0.3	1.0
1	0.7	1.0

b	ρ_{BD}	λ_{DB}
0	0.6	1.0
1	0.4	1.0

C				
c	ρ	λ	β	p(c)
0	0.425	1.0	0.425	0.425
1	0.575	1.0	0.575	0.575

D				
d	ρ	λ	β	p(d)
0	0.462	1.0	0.462	0.462
1	0.538	1.0	0.538	0.538

E				
e	ρ	λ	β	p(e)
0	0.1	1.0	0.1	0.1
1	0.9	1.0	0.9	0.9

d	ρ_{DF}	λ_{FD}
0	0.462	1.0
1	0.538	1.0

d	ρ_{DG}	λ_{GD}
0	0.462	1.0
1	0.538	1.0

e	ρ_{EG}	λ_{GE}
0	0.1	1.0
1	0.9	1.0

F				
f	ρ	λ	β	p(f)
0	0.531	1.0	0.531	0.531
1	0.469	1.0	0.469	0.469

G				
g	ρ	λ	β	p(g)
0	0.419	1.0	0.419	0.419
1	0.581	1.0	0.581	0.581

PROPAGATION IN CLUSTER TREES

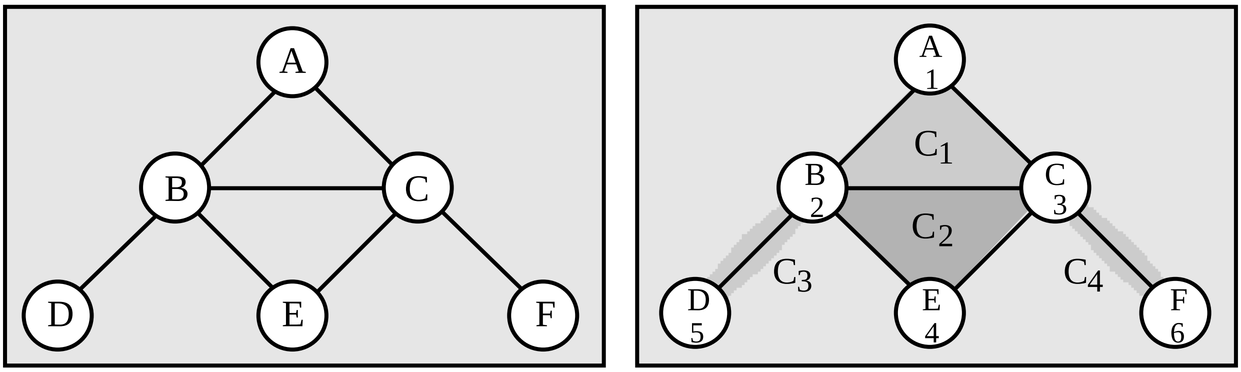
1. Absorb $E = e$ in the potential functions Ψ.
2. Obtain a chain of cliques $(C_1, \ldots, C_m)$ satisfying the running intersection property. For each clique C_i, choose as neighbor C_j, with $j < i$ and $S_i \subset C_j$.

Iteration Steps:

4. For $i = m$ to 1 (backwards) do
   (a) Compute $m_i(s_i) = \sum_{r_i} \psi_i(c_i)$.
   (b) Let $p(r_i|s_i) = \psi_i(c_i)/m_i(s_i)$.
   (c) Replace the potential function $\psi_j(c_j)$ of the neighboring clique C_j of clique C_i by $\psi_j(c_j) \leftarrow \psi_j(c_j)m_i(s_i)$.
5. Let $p(c_1) = p(r_1|s_1) = p(r_1)$.
6. For $i = 2$ to m (forwards) do
   (a) Compute $p(s_i)$ by marginalizing the JPD $p(c_j)$ of the neighboring clique of C_i, C_j.
   (b) Let $p(c_i) = p(r_i|s_i)p(s_i)$.
7. For $i = 1$ to n do
   (a) Choose the smallest clique C_j containing X_i.
   (b) Let $p(x_i|e) \propto \sum_{c_j \setminus x_i} p(c_j)$.

Cantabria
University

EXAMPLE OF PROPAGATION IN CLUSTER TREES



(a) (b)