Propagation in Belief Networks
Objective

**Given:** Belief network $(V, E, P)$ with tree structure and $P(V) > 0$. Set $W \subseteq V$ of instantiated variables where a priori knowledge $W \neq \emptyset$ is allowed.

**Desired:** $P(B | W)$ for all $B \in V$

**Notation:**
- $W_B^-$ subset of those variables of $W$ that belong to the subtree of $(V, E)$ that has root $B$
- $W_B^+ = W \setminus W_B^-$
- $s(B)$ set of direct successors of $B$
- $\Omega_B$ domain of $B$
- $b^*$ value that $B$ is instantiated with
Example

\[ W_B = \{ F, K, L, M \} \]

\[ W_B^+ = \{ F, K \} \]

\[ W_B^- = \{ L, M \} \]

\[ s(B) = \{ C, M, N \} \]
Decomposition in the Tree

\[
P(B = b \mid W) = \frac{P(b \mid W_B^- \cup W_B^+)}{P(W_B^- \cup W_B^+)} \quad \text{with } B \not\in W
\]

\[
= \frac{P(W_B^- \cup W_B^+ \cup \{b\})}{P(W_B^- \cup W_B^+)}
\]

\[
= \frac{P(W_B^- \cup W_B^+ \mid b)P(b)}{P(W_B^- \cup W_B^+)}
\]

\[
= \frac{P(W_B^- \mid b)P(W_B^+ \mid b)P(b)}{P(W_B^- \cup W_B^+)}
\]

\[
= \beta_{B,W} \underbrace{P(W_B^- \mid b)}_{\text{Evidence from “below”}} \underbrace{P(b \mid W_B^+)}_{\text{Evidence from “above”}}
\]
\[ \pi- \text{ and } \lambda-\text{Values} \]

Since we ignore the constant \( \beta_{B,W} \) for the derivations below, the following designations are used instead of \( P(\cdot) \):

\[ \text{\pi-values and \lambda-values} \]

Let \( B \in V \) be a variable and \( b \in \Omega_B \) a value of its domain. We define the \( \pi- \) and \( \lambda- \)values as follows:

\[
\lambda(b) = \begin{cases} 
    P(W_B^- | b) & \text{if } B \not\in W \\
    1 & \text{if } B \in W \land b^* = b \\
    0 & \text{if } B \in W \land b^* \neq b
\end{cases}
\]

\[
\pi(b) = P(b | W_B^+) 
\]
\( \pi - \) and \( \lambda - \)Values

\[
\lambda(b) = \prod_{C \in s(B)} P(W_C^- | b) \quad \text{if } B \in W
\]

\[
\lambda(b) = 1 \quad \text{if } B \text{ leaf in } (V, E)
\]

\[
\pi(b) = P(b) \quad \text{if } B \text{ root in } (V, E)
\]

\[
P(b | W) = \alpha_{B,W} \cdot \lambda(b) \cdot \pi(b)
\]
\textbf{\(\lambda\)-Message}

\textbf{\(\lambda\)-message}

Let \( B \in V \) be an attribute and \( C \in s(B) \) its direct children with the respective domains \( \text{dom}(B) = \{B_1, \ldots, b_i, \ldots, b_k\} \) and \( \text{dom}(C) = \{c_1, \ldots, c_j, \ldots, c_m\} \).

\[ \lambda_{C \to B}(b_i) \overset{\text{Def}}{=} \sum_{j=1}^{m} P(c_j \mid b_i) \cdot \lambda(c_j), \quad i = 1, \ldots, k \]

The vector

\[ \vec{\lambda}_{C \to B} \overset{\text{Def}}{=} \left( \lambda_{C \to B}(b_i) \right)_{i=1}^{k} \]

is called \textit{\(\lambda\)-message} from \( C \) to \( B \).
Let $B \in V$ an attribute an $b \in \text{dom}(B)$ a value of its domain. Then

$$
\lambda(b) = \begin{cases}
\rho_{B,W} \cdot \prod_{C \in s(B)} \lambda_C(b) & \text{if } B \notin W \\
1 & \text{if } B \in W \land b = b^* \\
0 & \text{if } B \in W \land b \neq b^*
\end{cases}
$$

with $\rho_{B,W}$ being a positive constant.
\(\pi\)-message

Let \(B \in V\) be a non-root node in \((V, E)\) and \(A \in V\) its parent with domain \(\text{dom}(A) = \{a_1, \ldots, a_j, \ldots, a_m\}\).

\[
\begin{align*}
\pi_A \rightarrow B(a_j) & \overset{\text{Def}}{=} \begin{cases}
\pi(a_j) \cdot \prod_{C \in s(A) \setminus \{B\}} \lambda_C(a_j) & \text{if } A \notin W \\
1 & \text{if } A \in W \land a = a^* \\
0 & \text{if } A \in W \land a \neq a^*
\end{cases}
\end{align*}
\]

The vector

\[
\vec{\pi}_A \rightarrow B \overset{\text{Def}}{=} \left(\pi_A \rightarrow B(a_j)\right)^m_{j=1}
\]

is called \(\pi\)-message from \(A\) to \(B\).
Let $B \in V$ be a non-root node in $(V, E)$ and $A$ the parent node of $B$. Further let $b \in \text{dom}(B)$ be a value of $B$’s domain.

$$\pi(b) = \mu_{B,W} \cdot \sum_{a \in \text{dom}(A)} P(b \mid a) \cdot \pi_{A\rightarrow B}(a)$$

Let $A \notin W$ a non-instantiated attribute and $P(V) > 0$.

$$\pi_{A\rightarrow B}(a_j) = \pi(a_j) \cdot \prod_{C \in s(A) \setminus \{B\}} \lambda_{C\rightarrow A}(a_j)$$

$$= \tau_{B,W} \cdot \frac{P(a_j \mid W)}{\lambda_{B\rightarrow A}(a_j)}$$
Propagation in Belief Trees

Belief Tree:

Parameters:

\[
\begin{align*}
P(a_1) &= 0.1 & P(b_1 | a_1) &= 0.7 \\
& & P(b_1 | a_2) &= 0.2 \\
& & P(d_1 | a_1) &= 0.8 & P(c_1 | b_1) &= 0.4 \\
& & P(d_1 | a_2) &= 0.4 & P(c_1 | b_2) &= 0.001 \\
\end{align*}
\]

Desired:

\[
\forall X \in \{A, B, C, D\} : P(X | \emptyset) = ?
\]
Belief Tree:

Initialization Phase:

Set all $\lambda$-messages and $\lambda$-values to 1.
Belief Tree:

<table>
<thead>
<tr>
<th></th>
<th>(\pi)</th>
<th>(\lambda)</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_1)</td>
<td>0.1</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>(a_2)</td>
<td>0.9</td>
<td>1</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Initialization Phase:

Set all \(\lambda\)-messages and \(\lambda\)-values to 1.

\[\pi(a_1) = P(a_1)\text{ and }\pi(a_2) = P(a_2)\]
Belief Tree:

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 0.1 & 1 & 0.1 \\
a_2 & 0.9 & 1 & 0.9 \\
\end{array}
\]

Initialization Phase:
Set all \(\lambda\)-messages and \(\lambda\)-values to 1.

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
d_1 & 1 & 1 \\
d_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
b_1 & 1 & 1 \\
b_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
c_1 & 1 & 1 \\
c_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 0.1 & 1 & 0.1 \\
a_2 & 0.9 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 1 & 1 & 0.1 \\
a_2 & 1 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
b_1 & 1 & 1 \\
b_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
c_1 & 1 & 1 \\
c_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 0.1 & 1 & 0.1 \\
a_2 & 0.9 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 1 & 1 & 0.1 \\
a_2 & 1 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
b_1 & 1 & 1 \\
b_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
c_1 & 1 & 1 \\
c_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 0.1 & 1 & 0.1 \\
a_2 & 0.9 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 1 & 1 & 0.1 \\
a_2 & 1 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
b_1 & 1 & 1 \\
b_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
c_1 & 1 & 1 \\
c_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 0.1 & 1 & 0.1 \\
a_2 & 0.9 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 1 & 1 & 0.1 \\
a_2 & 1 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
b_1 & 1 & 1 \\
b_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
c_1 & 1 & 1 \\
c_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 0.1 & 1 & 0.1 \\
a_2 & 0.9 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
a_1 & 1 & 1 & 0.1 \\
a_2 & 1 & 1 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
b_1 & 1 & 1 \\
b_2 & 1 & 1 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
\pi & \lambda & P \\
\hline
c_1 & 1 & 1 \\
c_2 & 1 & 1 \\
\end{array}
\]
Propagation in Belief Trees (5)

Belief Tree:

Initialization Phase:

Set all $\lambda$-messages and $\lambda$-values to 1.

$\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.

A sends $\pi$-messages to $B$ and $D$.

$B$ and $D$ update their $\pi$-values.
Belief Tree:

<table>
<thead>
<tr>
<th></th>
<th>$\pi$</th>
<th>$\lambda$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>0.1</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>$a_2$</td>
<td>0.9</td>
<td>1</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Initialization Phase:

- Set all $\lambda$-messages and $\lambda$-values to 1.
- $\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.
- $A$ sends $\pi$-messages to $B$ and $D$.
- $B$ and $D$ update their $\pi$-values.
- $B$ sends $\pi$-message to $C$. 

<table>
<thead>
<tr>
<th></th>
<th>$\pi$</th>
<th>$\lambda$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1$</td>
<td>0.44</td>
<td>1</td>
<td>0.44</td>
</tr>
<tr>
<td>$d_2$</td>
<td>0.56</td>
<td>1</td>
<td>0.56</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\pi$</th>
<th>$\lambda$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$</td>
<td>0.25</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>$b_2$</td>
<td>0.75</td>
<td>1</td>
<td>0.75</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>$\pi$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>$c_2$</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
Belief Tree:

```
<table>
<thead>
<tr>
<th></th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>0.1</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>a2</td>
<td>0.9</td>
<td>1</td>
<td>0.9</td>
</tr>
</tbody>
</table>
```

Initialization Phase:

Set all λ-messages and λ-values to 1.

- $\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.

A sends $\pi$-messages to $B$ and $D$.

$B$ and $D$ update their $\pi$-values.

$B$ sends $\pi$-message to $C$.

$C$ updates its $\pi$-value.
Propagation in Belief Trees (8)

**Belief Tree:**

**Initialization Phase:**

Set all $\lambda$-messages and $\lambda$-values to 1.

$\pi(a_1) = P(a_1)$ and $\pi(a_2) = P(a_2)$.

$A$ sends $\pi$-messages to $B$ and $D$.

$B$ and $D$ update their $\pi$-values.

$B$ sends $\pi$-message to $C$.

$C$ updates its $\pi$-value.

Initialization finished.
Larger Network (1): Parameters

\[
\begin{array}{c|cc}
P(A) & a_1 & a_2 \\
\hline
\emptyset & 0 & \\
a_1 & 0.4 & \\
a_2 & 0.6 & \\
\end{array}
\]

\[
\begin{array}{c|cc}
P(B \mid A) & a_1 & a_2 \\
\hline
b_1 & 0.2 & 0.3 \\
b_2 & 0.8 & 0.7 \\
\end{array}
\]

\[
\begin{array}{c|cc}
P(C \mid A) & a_1 & a_2 \\
\hline
c_1 & 0.1 & 0.25 \\
c_2 & 0.9 & 0.75 \\
\end{array}
\]

\[
\begin{array}{c|cc}
P(F \mid C) & c_1 & c_2 \\
\hline
f_1 & 0.3 & 0.6 \\
f_2 & 0.7 & 0.4 \\
\end{array}
\]

\[
\begin{array}{c|cc}
P(D \mid B) & b_1 & b_2 \\
\hline
d_1 & 0.5 & 0.35 \\
d_2 & 0.5 & 0.65 \\
\end{array}
\]

\[
\begin{array}{c|cc}
P(E \mid B) & b_1 & b_2 \\
\hline
e_1 & 0.15 & 0.45 \\
e_2 & 0.85 & 0.55 \\
\end{array}
\]

\[
\begin{array}{c|cc}
P(G \mid C) & c_1 & c_2 \\
\hline
g_1 & 0.25 & 0.1 \\
g_2 & 0.75 & 0.9 \\
\end{array}
\]

\[
\begin{array}{c|cc}
P(H \mid F) & f_1 & f_2 \\
\hline
h_1 & 0.65 & 0.2 \\
h_2 & 0.35 & 0.8 \\
\end{array}
\]

\[
\begin{array}{c|cc}
P(I \mid F) & f_1 & f_2 \\
\hline
i_1 & 0.25 & 0.5 \\
i_2 & 0.75 & 0.5 \\
\end{array}
\]
Larger Network (2): After Initialization

A

B

C

D

E

F

G

H

I

<table>
<thead>
<tr>
<th>A</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>a₁</td>
<td>0.4</td>
<td>1</td>
<td>0.4</td>
</tr>
<tr>
<td>a₂</td>
<td>0.6</td>
<td>1</td>
<td>0.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>b₁</td>
<td>0.26</td>
<td>1</td>
<td>0.26</td>
</tr>
<tr>
<td>b₂</td>
<td>0.74</td>
<td>1</td>
<td>0.74</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>c₁</td>
<td>0.19</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>c₂</td>
<td>0.81</td>
<td>1</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>d₁</td>
<td>0.389</td>
<td>1</td>
<td>0.389</td>
</tr>
<tr>
<td>d₂</td>
<td>0.611</td>
<td>1</td>
<td>0.611</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>E</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>e₁</td>
<td>0.372</td>
<td>1</td>
<td>0.327</td>
</tr>
<tr>
<td>e₂</td>
<td>0.628</td>
<td>1</td>
<td>0.628</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>f₁</td>
<td>0.543</td>
<td>1</td>
<td>0.543</td>
</tr>
<tr>
<td>f₂</td>
<td>0.457</td>
<td>1</td>
<td>0.457</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>G</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>g₁</td>
<td>0.1285</td>
<td>1</td>
<td>0.1285</td>
</tr>
<tr>
<td>g₂</td>
<td>0.8715</td>
<td>1</td>
<td>0.8715</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>H</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h₁</td>
<td>0.4444</td>
<td>1</td>
<td>0.4444</td>
</tr>
<tr>
<td>h₂</td>
<td>0.5556</td>
<td>1</td>
<td>0.5556</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>I</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>i₁</td>
<td>0.3643</td>
<td>1</td>
<td>0.3643</td>
</tr>
<tr>
<td>i₂</td>
<td>0.6357</td>
<td>1</td>
<td>0.6357</td>
</tr>
</tbody>
</table>
Larger Network (3): Set Evidence $e_1, g_1, h_1$
Larger Network (4): Propagate Evidence

- A
  - \(a_1\) 0.4 1 0.4
  - \(a_2\) 0.6 1 0.6

- B
  - \(b_1\) 0.26 1 0.26
  - \(b_2\) 0.74 1 0.74

- C
  - \(c_1\) 0.19 1 0.19
  - \(c_2\) 0.81 1 0.81

- D
  - \(d_1\) 0.389 1 0.389
  - \(d_2\) 0.611 1 0.611

- E
  - \(e_1\) 1 1
  - \(e_2\) 0 0

- F
  - \(f_1\) 0.543 1 0.543
  - \(f_2\) 0.457 1 0.457

- G
  - \(g_1\) 1 1
  - \(g_2\) 0 0

- H
  - \(h_1\) 1 1
  - \(h_2\) 0 0

- I
  - \(i_1\) 0.3643 1 0.3643
  - \(i_2\) 0.6357 1 0.6357
Larger Network (5): Propagate Evidence, cont.

\[ A \]
<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( \lambda )</th>
<th>( P )</th>
</tr>
</thead>
</table>
| \( a_1 \) | 0.4 | 1 
| \( a_2 \) | 0.6 | 1 

\[ B \]
<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( \lambda )</th>
<th>( P )</th>
</tr>
</thead>
</table>
| \( b_1 \) | 0.26 | 0.15 
| \( b_2 \) | 0.74 | 0.45 

\[ C \]
<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( \lambda )</th>
<th>( P )</th>
</tr>
</thead>
</table>
| \( c_1 \) | 0.19 | 0.25 
| \( c_2 \) | 0.81 | 0.1 

\[ D \]
<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( \lambda )</th>
<th>( P )</th>
</tr>
</thead>
</table>
| \( d_1 \) | 0.389 | 1 
| \( d_2 \) | 0.611 | 1 

\[ E \]
<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( \lambda )</th>
<th>( P )</th>
</tr>
</thead>
</table>
| \( e_1 \) | 1 | 1 
| \( e_2 \) | 0 | 0 

\[ F \]
<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( \lambda )</th>
<th>( P )</th>
</tr>
</thead>
</table>
| \( f_1 \) | 0.543 | 0.65 
| \( f_2 \) | 0.457 | 0.2 

\[ G \]
<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( \lambda )</th>
<th>( P )</th>
</tr>
</thead>
</table>
| \( g_1 \) | 1 | 1 
| \( g_2 \) | 0 | 0 

\[ H \]
<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( \lambda )</th>
<th>( P )</th>
</tr>
</thead>
</table>
| \( h_1 \) | 1 | 1 
| \( h_2 \) | 0 | 0 

\[ I \]
<table>
<thead>
<tr>
<th>( \pi )</th>
<th>( \lambda )</th>
<th>( P )</th>
</tr>
</thead>
</table>
| \( i_1 \) | 0.3643 | 1 
| \( i_2 \) | 0.6357 | 1 

Rudolf Kruse, Matthias Steinbrecher, Pascal Held
Bayesian Networks

\[ A \quad \pi \quad \lambda \quad P \]
\[ a_1 \quad 0.4 \quad 1 \quad 0.4 \]
\[ a_2 \quad 0.6 \quad 1 \quad 0.6 \]

\[ B \quad \pi \quad \lambda \quad P \]
\[ b_1 \quad 0.26 \quad 0.15 \quad 0.1048 \]
\[ b_2 \quad 0.74 \quad 0.45 \quad 0.8952 \]

\[ (a_1, 0.39, 0.36) \]

\[ C \quad \pi \quad \lambda \quad P \]
\[ c_1 \quad 0.19 \quad 0.25 \quad 0.3696 \]
\[ c_2 \quad 0.81 \quad 0.1 \quad 0.6304 \]

\[ (a_1, 0.335, 0.47) \]

\[ D \quad \pi \quad \lambda \quad P \]
\[ d_1 \quad 0.389 \quad 1 \quad 0.389 \]
\[ d_2 \quad 0.611 \quad 1 \quad 0.611 \]

\[ (b_1, 0.1048, 0.8952) \]

\[ E \quad \pi \quad \lambda \quad P \]
\[ e_1 \quad 1 \quad 1 \]
\[ e_2 \quad 0 \quad 0 \]

\[ (b_1, 0.39, 0.36) \]

\[ F \quad \pi \quad \lambda \quad P \]
\[ f_1 \quad 0.543 \quad 0.65 \quad 0.7943 \]
\[ f_2 \quad 0.457 \quad 0.2 \quad 0.2057 \]

\[ (c_1, 0.335, 0.47) \]

\[ (f_1, 0.7943, 0.2057) \]

\[ G \quad \pi \quad \lambda \quad P \]
\[ g_1 \quad 1 \quad 1 \]
\[ g_2 \quad 0 \quad 0 \]

\[ (c_1, 0.335, 0.47) \]

\[ (f_1, 0.7943, 0.2057) \]

\[ H \quad \pi \quad \lambda \quad P \]
\[ h_1 \quad 1 \quad 1 \]
\[ h_2 \quad 0 \quad 0 \]

\[ (f_1, 0.7943, 0.2057) \]

\[ (h_1, 1, 1) \]

\[ (h_2, 0, 0) \]

\[ (i_1, 0.3643, 1) \]

\[ (i_2, 0.6357, 1) \]

Rudolf Kruse, Matthias Steinbrecher, Pascal Held

Bayesian Networks
Larger Network (7): Propagate Evidence, cont.

\[
\begin{array}{cccc}
A & \pi & \lambda & P \\
0.4 & 0.39 & 0.4194 \\
0.6 & 0.36 & 0.5806 \\
\end{array}
\]

\[
\begin{array}{cccc}
B & \pi & \lambda & P \\
b_1 & 0.26 & 0.15 & 0.1048 \\
b_2 & 0.74 & 0.45 & 0.8952 \\
\end{array}
\]

\[
\begin{array}{cccc}
D & \pi & \lambda & P \\
d_1 & 0.3657 & 1 & 0.3657 \\
d_2 & 0.6343 & 1 & 0.6343 \\
\end{array}
\]

\[
\begin{array}{cccc}
E & \pi & \lambda & P \\
e_1 & 1 & 1 & \\
e_2 & 0 & 0 & \\
\end{array}
\]

\[
\begin{array}{cccc}
F & \pi & \lambda & P \\
f_1 & 0.543 & 0.65 & 0.7943 \\
f_2 & 0.457 & 0.2 & 0.2057 \\
\end{array}
\]

\[
\begin{array}{cccc}
C & \pi & \lambda & P \\
c_1 & 0.19 & 0.0838 & 0.2948 \\
c_2 & 0.81 & 0.047 & 0.7052 \\
\end{array}
\]

\[
\begin{array}{cccc}
G & \pi & \lambda & P \\
g_1 & 1 & 1 & \\
g_2 & 0 & 0 & \\
\end{array}
\]

\[
\begin{array}{cccc}
I & \pi & \lambda & P \\
i_1 & 0.3014 & 1 & 0.3014 \\
i_2 & 0.6986 & 1 & 0.6986 \\
\end{array}
\]
Larger Network (8): Propagate Evidence, cont.

\[ A \quad \pi \quad \lambda \quad P \]
\[ a_1 \quad 0.4 \quad 0.39 \quad 0.4194 \]
\[ a_2 \quad 0.6 \quad 0.36 \quad 0.5806 \]

\[ B \quad \pi \quad \lambda \quad P \]
\[ b_1 \quad 0.26 \quad 0.15 \quad 0.1048 \]
\[ b_2 \quad 0.74 \quad 0.45 \quad 0.8952 \]

\[ C \quad \pi \quad \lambda \quad P \]
\[ c_1 \quad 0.19 \quad 0.0838 \quad 0.2948 \]
\[ c_2 \quad 0.81 \quad 0.047 \quad 0.7052 \]

\[ D \quad \pi \quad \lambda \quad P \]
\[ d_1 \quad 0.3657 \quad 1 \quad 0.3657 \]
\[ d_2 \quad 0.6343 \quad 1 \quad 0.6343 \]

\[ E \quad \pi \quad \lambda \quad P \]
\[ e_1 \quad 1 \quad 1 \]
\[ e_2 \quad 0 \quad 0 \]

\[ F \quad \pi \quad \lambda \quad P \]
\[ f_1 \quad 0.543 \quad 0.65 \quad 0.7943 \]
\[ f_2 \quad 0.457 \quad 0.2 \quad 0.2057 \]

\[ G \quad \pi \quad \lambda \quad P \]
\[ g_1 \quad 1 \quad 1 \]
\[ g_2 \quad 0 \quad 0 \]

\[ H \quad \pi \quad \lambda \quad P \]
\[ h_1 \quad 1 \quad 1 \]
\[ h_2 \quad 0 \quad 0 \]

\[ I \quad \pi \quad \lambda \quad P \]
\[ i_1 \quad 0.3014 \quad 1 \quad 0.3014 \]
\[ i_2 \quad 0.6986 \quad 1 \quad 0.6986 \]
Larger Network (9): Propagate Evidence, cont.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>0.4</td>
<td>0.0198</td>
<td>0.3945</td>
<td></td>
</tr>
<tr>
<td>a2</td>
<td>0.6</td>
<td>0.0202</td>
<td>0.6055</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>b1</td>
<td>0.26</td>
<td>0.15</td>
<td>0.1048</td>
<td></td>
</tr>
<tr>
<td>b2</td>
<td>0.74</td>
<td>0.45</td>
<td>0.8952</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>c1</td>
<td>0.19</td>
<td>0.0838</td>
<td>0.2948</td>
<td></td>
</tr>
<tr>
<td>c2</td>
<td>0.81</td>
<td>0.047</td>
<td>0.7052</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>d1</td>
<td>0.3657</td>
<td>1</td>
<td>0.3657</td>
<td></td>
</tr>
<tr>
<td>d2</td>
<td>0.6343</td>
<td>1</td>
<td>0.6343</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>E</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>e1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>e2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>f1</td>
<td>0.4891</td>
<td>0.65</td>
<td>0.7508</td>
<td></td>
</tr>
<tr>
<td>f2</td>
<td>0.5109</td>
<td>0.2</td>
<td>0.2432</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>G</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>g1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>g2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>H</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>h1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>h2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>π</th>
<th>λ</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>i1</td>
<td>0.3014</td>
<td>1</td>
<td>0.3014</td>
<td></td>
</tr>
<tr>
<td>i2</td>
<td>0.6986</td>
<td>1</td>
<td>0.6986</td>
<td></td>
</tr>
</tbody>
</table>
Larger Network (10): Propagate Evidence, cont.

\[
\begin{array}{cccc}
A & \pi & \lambda & P \\
\hline
a_1 & 0.4 & 0.0198 & 0.3945 \\
a_2 & 0.6 & 0.0202 & 0.6055 \\
\end{array}
\]

\[
\begin{array}{cccc}
B & \pi & \lambda & P \\
\hline
b_1 & 0.26 & 0.15 & 0.1048 \\
b_2 & 0.74 & 0.45 & 0.8952 \\
\end{array}
\]

\[
\begin{array}{cccc}
C & \pi & \lambda & P \\
\hline
c_1 & 0.19 & 0.0838 & 0.2948 \\
c_2 & 0.81 & 0.047 & 0.7052 \\
\end{array}
\]

\[
\begin{array}{cccc}
D & \pi & \lambda & P \\
\hline
d_1 & 0.3657 & 1 & 0.3657 \\
d_2 & 0.6343 & 1 & 0.6343 \\
\end{array}
\]

\[
\begin{array}{cccc}
E & \pi & \lambda & P \\
\hline
e_1 & 1 & 1 & \\
e_2 & 0 & 0 & \\
\end{array}
\]

\[
\begin{array}{cccc}
F & \pi & \lambda & P \\
\hline
f_1 & 0.4891 & 0.65 & 0.7508 \\
f_2 & 0.5109 & 0.2 & 0.2432 \\
\end{array}
\]

\[
\begin{array}{cccc}
G & \pi & \lambda & P \\
\hline
g_1 & 1 & 1 & \\
g_2 & 0 & 0 & \\
\end{array}
\]

\[
\begin{array}{cccc}
H & \pi & \lambda & P \\
\hline
h_1 & 1 & 1 & \\
h_2 & 0 & 0 & \\
\end{array}
\]

\[
\begin{array}{cccc}
I & \pi & \lambda & P \\
\hline
i_1 & 0.3014 & 1 & 0.3014 \\
i_2 & 0.6986 & 1 & 0.6986 \\
\end{array}
\]

\[
\begin{array}{cccc}
A & \pi & \lambda & P \\
\hline
a_1 & 0.4 & 0.0198 & 0.3945 \\
a_2 & 0.6 & 0.0202 & 0.6055 \\
\end{array}
\]

\[
\begin{array}{cccc}
B & \pi & \lambda & P \\
\hline
b_1 & 0.26 & 0.15 & 0.1048 \\
b_2 & 0.74 & 0.45 & 0.8952 \\
\end{array}
\]

\[
\begin{array}{cccc}
C & \pi & \lambda & P \\
\hline
c_1 & 0.19 & 0.0838 & 0.2948 \\
c_2 & 0.81 & 0.047 & 0.7052 \\
\end{array}
\]

\[
\begin{array}{cccc}
D & \pi & \lambda & P \\
\hline
d_1 & 0.3657 & 1 & 0.3657 \\
d_2 & 0.6343 & 1 & 0.6343 \\
\end{array}
\]

\[
\begin{array}{cccc}
E & \pi & \lambda & P \\
\hline
e_1 & 1 & 1 & \\
e_2 & 0 & 0 & \\
\end{array}
\]

\[
\begin{array}{cccc}
F & \pi & \lambda & P \\
\hline
f_1 & 0.4891 & 0.65 & 0.7508 \\
f_2 & 0.5109 & 0.2 & 0.2432 \\
\end{array}
\]

\[
\begin{array}{cccc}
G & \pi & \lambda & P \\
\hline
g_1 & 1 & 1 & \\
g_2 & 0 & 0 & \\
\end{array}
\]

\[
\begin{array}{cccc}
H & \pi & \lambda & P \\
\hline
h_1 & 1 & 1 & \\
h_2 & 0 & 0 & \\
\end{array}
\]

\[
\begin{array}{cccc}
I & \pi & \lambda & P \\
\hline
i_1 & 0.3014 & 1 & 0.3014 \\
i_2 & 0.6986 & 1 & 0.6986 \\
\end{array}
\]

Rudolf Kruse, Matthias Steinbrecher, Pascal Held

Bayesian Networks
Larger Network (12): Propagate Evidence, cont.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{4}{|c|}{A, \pi, \lambda, P} \\
\hline
a_1 & 0.4 & 0.0198 & 0.3945 \\
\hline
a_2 & 0.6 & 0.0202 & 0.6055 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{4}{|c|}{B, \pi, \lambda, P} \\
\hline
b_1 & 0.7077 & 0.15 & 0.1061 \\
\hline
b_2 & 1.9865 & 0.45 & 0.8939 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{4}{|c|}{C, \pi, \lambda, P} \\
\hline
c_1 & 0.1871 & 0.0838 & 0.2910 \\
\hline
c_2 & 0.8129 & 0.047 & 0.7090 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{4}{|c|}{D, \pi, \lambda, P} \\
\hline
d_1 & 0.3657 & 1 & 0.3657 \\
\hline
d_2 & 0.6343 & 1 & 0.6343 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{4}{|c|}{E, \pi, \lambda, P} \\
\hline
e_1 & 1 & 1 & \ \\
\hline
e_2 & 0 & 0 & \ \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{4}{|c|}{F, \pi, \lambda, P} \\
\hline
f_1 & 0.4891 & 0.65 & 0.7508 \\
\hline
f_2 & 0.5109 & 0.2 & 0.2432 \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{4}{|c|}{G, \pi, \lambda, P} \\
\hline
g_1 & 1 & 1 & \ \\
\hline
g_2 & 0 & 0 & \ \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{4}{|c|}{H, \pi, \lambda, P} \\
\hline
h_1 & 1 & 1 & \ \\
\hline
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{4}{|c|}{I, \pi, \lambda, P} \\
\hline
i_1 & 0.3108 & 1 & 0.3108 \\
\hline
i_2 & 0.6892 & 1 & 0.6892 \\
\hline
\end{tabular}
\end{table}
Larger Network (13): Propagate Evidence, cont.

\[
\begin{array}{cccc}
A & \pi & \lambda & P \\
\hline
a_1 & 0.4 & 0.0198 & 0.3945 \\
a_2 & 0.6 & 0.0202 & 0.6055 \\
\end{array}
\]

\[
\begin{array}{cccc}
B & \pi & \lambda & P \\
\hline
b_1 & 0.7077 & 0.15 & 0.1061 \\
b_2 & 1.9865 & 0.45 & 0.8939 \\
\end{array}
\]

\[
\begin{array}{cccc}
C & \pi & \lambda & P \\
\hline
\end{array}
\]

\[
\begin{array}{cccc}
D & \pi & \lambda & P \\
\hline
d_1 & 0.3659 & 1 & 0.3659 \\
d_2 & 0.6341 & 1 & 0.6341 \\
\end{array}
\]

\[
\begin{array}{cccc}
E & \pi & \lambda & P \\
\hline
e_1 & 1 & 1 & 0 \\
e_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
F & \pi & \lambda & P \\
\hline
f_1 & 1.1657 & 0.65 & 0.7577 \\
f_2 & 1.2115 & 0.2 & 0.2423 \\
\end{array}
\]

\[
\begin{array}{cccc}
G & \pi & \lambda & P \\
\hline
g_1 & 1 & 1 & 0 \\
g_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
H & \pi & \lambda & P \\
\hline
h_1 & 1 & 1 & 0 \\
h_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
I & \pi & \lambda & P \\
\hline
i_1 & 0.3108 & 1 & 0.3108 \\
i_2 & 0.6892 & 1 & 0.6892 \\
\end{array}
\]

\[
\begin{array}{cccc}
J & \pi & \lambda & P \\
\hline
j_1 & 1.1657 & 0.65 & 0.7577 \\
j_2 & 1.2115 & 0.2 & 0.2423 \\
\end{array}
\]

\[
\begin{array}{cccc}
K & \pi & \lambda & P \\
\hline
k_1 & 1 & 1 & 0 \\
k_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
L & \pi & \lambda & P \\
\hline
l_1 & 0.3659 & 1 & 0.3659 \\
l_2 & 0.6341 & 1 & 0.6341 \\
\end{array}
\]

\[
\begin{array}{cccc}
M & \pi & \lambda & P \\
\hline
m_1 & 1 & 1 & 0 \\
m_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
N & \pi & \lambda & P \\
\hline
n_1 & 0.3108 & 1 & 0.3108 \\
n_2 & 0.6892 & 1 & 0.6892 \\
\end{array}
\]

\[
\begin{array}{cccc}
O & \pi & \lambda & P \\
\hline
o_1 & 1 & 1 & 0 \\
o_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
P & \pi & \lambda & P \\
\hline
p_1 & 0.7077 & 0.15 & 0.1061 \\
p_2 & 1.9865 & 0.45 & 0.8939 \\
\end{array}
\]

\[
\begin{array}{cccc}
Q & \pi & \lambda & P \\
\hline
q_1 & 1 & 1 & 0 \\
q_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
R & \pi & \lambda & P \\
\hline
r_1 & 0.3659 & 1 & 0.3659 \\
r_2 & 0.6341 & 1 & 0.6341 \\
\end{array}
\]

\[
\begin{array}{cccc}
S & \pi & \lambda & P \\
\hline
s_1 & 1 & 1 & 0 \\
s_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
T & \pi & \lambda & P \\
\hline
t_1 & 0.3108 & 1 & 0.3108 \\
t_2 & 0.6892 & 1 & 0.6892 \\
\end{array}
\]

\[
\begin{array}{cccc}
U & \pi & \lambda & P \\
\hline
u_1 & 1 & 1 & 0 \\
u_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
V & \pi & \lambda & P \\
\hline
v_1 & 0.3659 & 1 & 0.3659 \\
v_2 & 0.6341 & 1 & 0.6341 \\
\end{array}
\]

\[
\begin{array}{cccc}
W & \pi & \lambda & P \\
\hline
w_1 & 1 & 1 & 0 \\
w_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
X & \pi & \lambda & P \\
\hline
x_1 & 0.3108 & 1 & 0.3108 \\
x_2 & 0.6892 & 1 & 0.6892 \\
\end{array}
\]

\[
\begin{array}{cccc}
Y & \pi & \lambda & P \\
\hline
y_1 & 1 & 1 & 0 \\
y_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{cccc}
Z & \pi & \lambda & P \\
\hline
z_1 & 0.3659 & 1 & 0.3659 \\
z_2 & 0.6341 & 1 & 0.6341 \\
\end{array}
\]
Larger Network (15): Finished

\[
\begin{array}{c|c|c|c}
A & \pi & \lambda & P \\
\hline
a_1 & 0.4 & 0.0198 & 0.3945 \\
a_2 & 0.6 & 0.0202 & 0.6055 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
B & \pi & \lambda & P \\
\hline
b_1 & 0.7077 & 0.15 & 0.1061 \\
b_2 & 1.9865 & 0.45 & 0.8939 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
C & \pi & \lambda & P \\
\hline
c_1 & 0.1871 & 0.0838 & 0.2910 \\
c_2 & 0.8129 & 0.047 & 0.7090 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
D & \pi & \lambda & P \\
\hline
d_1 & 0.3659 & 1 & 0.3659 \\
d_2 & 0.6341 & 1 & 0.6341 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
E & \pi & \lambda & P \\
\hline
e_1 & 1 & 1 & 1 \\
e_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
F & \pi & \lambda & P \\
\hline
f_1 & 1.1657 & 0.65 & 0.7577 \\
f_2 & 1.2115 & 0.2 & 0.2423 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
G & \pi & \lambda & P \\
\hline
g_1 & 1 & 1 & 0 \\
g_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
H & \pi & \lambda & P \\
\hline
h_1 & 1 & 1 & 1 \\
h_2 & 0 & 0 & 0 \\
\end{array}
\]

\[
\begin{array}{c|c|c|c}
I & \pi & \lambda & P \\
\hline
i_1 & 0.3106 & 1 & 0.3106 \\
i_2 & 0.6894 & 1 & 0.6894 \\
\end{array}
\]
Propagation in Clique Trees
Problems

The propagation algorithm as presented can only deal with *trees*. Can be extended to *polytrees* (i.e. singly connected graphs with multiple parents per node).

However, it cannot handle networks that contain loops!
Main Objectives:

Transform the cyclic directed graph into a secondary structure without cycles.

Find a decomposition of the underlying joint distribution.

Task:

Combine nodes of the original (primary) graph structure.

These groups form the nodes of a secondary structure.

Find a transformation that yields tree structure.
Secondary Structure:

We will generate an undirected graph mimicking (some of) the conditional independence statements of the cyclic directed graph.

Maximal cliques are identified and form the nodes of the secondary structure.

Specify a so-called potential function for every clique such that the product of all potentials yields the initial joint distribution.

In order to propagate evidence, create a tree from the clique nodes such that the following property is satisfied:

If two cliques have some attributes in common, then these attributes have to be contained in every clique of the path connecting the two cliques. (called the running intersection property, RIP)

Justification:

Tree: Unique path of evidence propagation.

RIP: Update of an attribute reaches all cliques which contain it.
Complete Graph

An undirected Graph $G = (V, E)$ is called complete, if every pair of (distinct) nodes is connected by an edge.

Induced Subgraph

Let $G = (V, E)$ be an undirected graph and $W \subseteq V$ a selection of nodes. Then, $G_W = (W, E_W)$ is called the subgraph of $G$ induced by $W$ with $E_W$ being

$$E_W = \{(u, v) \in E \mid u, v \in W\}.$$
Complete Set, Clique

Let $G = (V, E)$ be an undirected graph. A set $W \subseteq V$ is called \textit{complete} iff it induces a complete subgraph. It is further called a \textit{clique}, iff $W$ is maximal, i.e. it is not possible to add a node to $W$ without violating the completeness condition.

a) $W$ is complete $\iff$ $W$ induces a complete subgraph

b) $W$ is a clique $\iff$ $W$ is complete and maximal

$C_1 = \{A, B, C, D\}$
$C_2 = \{B, D, E\}$
$C_3 = \{E, F\}$

3 cliques
Perfect Ordering

Let \( G = (V, E) \) be an undirected graph with \( n \) nodes and \( \alpha = \langle v_1, \ldots, v_n \rangle \) a total ordering on \( V \). Then, \( \alpha \) is called \textit{perfect}, if the following sets

\[
\text{adj}(v_i) \cap \{v_1, \ldots, v_{i-1}\} \quad i = 1, \ldots, n
\]

are complete, where \( \text{adj}(v_i) = \{w \mid (v_i, w) \in E\} \) returns the adjacent nodes of \( v_i \).

\[
\begin{array}{c|c|c|c}
 i & \text{adj}(v_i) & \text{adj}(v_i) \cap \{v_1, \ldots, v_{i-1}\} & \text{complete} \\
1 & \{C\} & \{C\} \cap \emptyset & \emptyset \quad \text{complete} \\
2 & \{A, D, F\} & \{A\} \cap \{A, D, F\} & \{A\} \quad \text{complete} \\
3 & \{C, B, E, F\} & \{A, C\} \cap \{C, B, E, F\} & \{C\} \quad \text{complete} \\
4 & \{G, C, D, E, H\} & \{A, C, D\} \cap \{G, C, D, E, H\} & \{C, D\} \quad \text{complete} \\
5 & \{B, D, F, H\} & \{A, C, D, F\} \cap \{B, D, F, H\} & \{D, F\} \quad \text{complete} \\
6 & \{D, E\} & \{A, C, D, F, E\} \cap \{D, E\} & \{D, E\} \quad \text{complete} \\
7 & \{F, E\} & \{A, C, D, F, E, B\} \cap \{F, E\} & \{F, E\} \quad \text{complete} \\
8 & \{F\} & \{A, C, D, F, E, B, H\} \cap \{F\} & \{F\} \quad \text{complete}
\end{array}
\]

\( \alpha = \langle A, C, D, F, E, B, H, G \rangle \)

\( \alpha \) is a perfect ordering
Running Intersection Property

Let $G = (V, E)$ be an undirected graph with $p$ cliques. An ordering of these cliques has the running intersection property (RIP), if for every $j > 1$ there exists an $i < j$ such that:

$$C_j \cap \left(C_1 \cup \cdots \cup C_{j-1}\right) \subseteq C_i$$

$$\xi = \langle C_1, C_2, C_3, C_4, C_5, C_6 \rangle$$

$\xi$ has running intersection property
If a node ordering $\alpha$ of an undirected graph $G = (V, E)$ is perfect and the cliques of $G$ are ordered according to the highest rank (w.r.t. $\alpha$) of the containing nodes, then this clique ordering has RIP.

<table>
<thead>
<tr>
<th>Clique</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>${A, C}$</td>
<td>$\max{\alpha(A), \alpha(C)}$</td>
</tr>
<tr>
<td>${C, D, F}$</td>
<td>$\max{\alpha(C), \alpha(D), \alpha(F)}$</td>
</tr>
<tr>
<td>${D, E, F}$</td>
<td>$\max{\alpha(D), \alpha(E), \alpha(F)}$</td>
</tr>
<tr>
<td>${B, D, E}$</td>
<td>$\max{\alpha(B), \alpha(D), \alpha(E)}$</td>
</tr>
<tr>
<td>${F, E, H}$</td>
<td>$\max{\alpha(F), \alpha(E), \alpha(H)}$</td>
</tr>
<tr>
<td>${F, G}$</td>
<td>$\max{\alpha(F), \alpha(G)}$</td>
</tr>
</tbody>
</table>

How to get a perfect ordering?
**Triangulated Graph**

An undirected graph is called *triangulated* if every simple loop (i.e. path with identical start and end node but with any other node occurring at most once) of length greater than 3 has a chord.
Maximum Cardinality Search

Let $G = (V, E)$ be an undirected graph. An ordering according *maximum cardinality search (MCS)* is obtained by first assigning 1 to an arbitrary node. If $n$ numbers are assigned the node that is connected to most of the nodes already numbered gets assigned number $n + 1$.

3 can be assigned to $D$ or $F$

6 can be assigned to $H$ or $B
An undirected graph is triangulated iff the ordering obtained by MCS is perfect.

To check whether a graph is triangulated is efficient to implement. The optimization problem that is related to the triangulation task is NP-hard. However, there are good heuristics.

**Moral Graph** (Repetition)

Let $G = (V, E)$ be a directed acyclic graph. If $u, w \in W$ are parents of $v \in V$ connect $u$ and $w$ with an (arbitrarily oriented) edge. After the removal of all edge directions the resulting graph $G_m = (V, E')$ is called the *moral graph* of $G$. 
Given directed graph.
Join-Tree Construction (2)

• Moral graph
Join-Tree Construction (3)

- Moral graph
- Triangulated graph
Join-Tree Construction (4)

- Moral graph
- Triangulated graph
- MCS yields perfect ordering
Join-Tree Construction (5)

- Moral graph
- Triangulated graph
- MCS yields perfect ordering
- Clique order has RIP
Join-Tree Construction (6)

- Moral graph
- Triangulated graph
- MCS yields perfect ordering
- Clique order has RIP
- Form a join-tree

Two cliques can be connected if they have a non-empty intersection. The generation of the tree follows the RIP. In case of a tie, connect cliques with the largest intersection. (e.g. $DBE—FED$ instead of $DBE—CFD$) Break remaining ties arbitrarily.
Example: Expert Knowledge

**Qualitative knowledge:**
Metastatic cancer is a possible cause of brain tumor, and is also an explanation for increased total serum calcium. In turn, either of these could explain a patient falling into a coma. Severe headache is also possibly associated with a brain tumor.

**Special case:**
The patient has heavy headache.

**Query:**
Will the patient fall into coma?
## Example: Choice of State Space

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Possible Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$ metastatic cancer</td>
<td>$\text{dom}(A) = {a_1, a_2} \cdot 1 = \text{existing}$</td>
</tr>
<tr>
<td>$B$ increased total serum calcium</td>
<td>$\text{dom}(B) = {b_1, b_2} \cdot 2 = \text{notexisting}$</td>
</tr>
<tr>
<td>$C$ brain tumor</td>
<td>$\text{dom}(C) = {c_1, c_2}$</td>
</tr>
<tr>
<td>$D$ coma</td>
<td>$\text{dom}(D) = {d_1, d_2}$</td>
</tr>
<tr>
<td>$E$ severe headache</td>
<td>$\text{dom}(E) = {e_1, e_2}$</td>
</tr>
</tbody>
</table>

Exhaustive state space:

$$\Omega = \text{dom}(A) \times \text{dom}(B) \times \text{dom}(C) \times \text{dom}(D) \times \text{dom}(E)$$

Marginal and conditional probabilities have to be specified!
Example: Qualitative Knowledge

\[
P(e_1 \mid c_1) = 0.8 \quad \}
\text{headaches common, but more common if tumor present}
\]
\[
P(e_1 \mid c_2) = 0.6 \quad \}
\text{headaches common, but more common if tumor present}
\]
\[
P(d_1 \mid b_1, c_1) = 0.8
\]
\[
P(d_1 \mid b_1, c_2) = 0.8
\]
\[
P(d_1 \mid b_1, c_2) = 0.8
\]
\[
P(d_1 \mid b_2, c_1) = 0.8
\]
\[
P(d_1 \mid b_2, c_2) = 0.05
\]
\[
P(b_1 \mid a_1) = 0.8 \quad \}
\text{increased calcium uncommon,}
\]
\[
P(b_1 \mid a_2) = 0.2 \quad \}
\text{but common consequence of metastases}
\]
\[
P(c_1 \mid a_1) = 0.2 \quad \}
\text{brain tumor rare, and uncommon consequence of metastases}
\]
\[
P(c_1 \mid a_2) = 0.05
\]
\[
P(a_1) = 0.2 \quad \}
\text{incidence of metastatic cancer in relevant clinic}
Propagation on Cliques (1)

Example: Metastatic Cancer

Dependencies

Moralization/Triangulation

MCS, hyper graph

Clique tree with separator sets
Quantitative knowledge:

<table>
<thead>
<tr>
<th>((a, b, c))</th>
<th>(P(a, b, c))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_1, b_1, c_1)</td>
<td>0.032</td>
</tr>
<tr>
<td>(a_2, b_1, c_1)</td>
<td>0.008</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>(a_2, b_2, c_2)</td>
<td>0.608</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>((b, c, d))</th>
<th>(P(b, c, d))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(b_1, c_1, d_1)</td>
<td>0.032</td>
</tr>
<tr>
<td>(b_2, c_1, d_1)</td>
<td>0.032</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>(b_2, c_2, d_2)</td>
<td>0.608</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>((c, e))</th>
<th>(P(c, e))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(c_1, e_1)</td>
<td>0.064</td>
</tr>
<tr>
<td>(c_2, e_1)</td>
<td>0.552</td>
</tr>
<tr>
<td>(c_1, e_2)</td>
<td>0.016</td>
</tr>
<tr>
<td>(c_2, e_2)</td>
<td>0.368</td>
</tr>
</tbody>
</table>

Potential representation:

\[
P(A, B, C, D, E, \ldots) = \frac{P(A | \emptyset) P(B | A) P(C | A) P(D | B C) P(E | C)}{P(A, B, C) P(B, C, D), P(C, E)}
\]

\[
= \frac{P(A, B, C) P(B, C, D), P(C, E)}{P(B C) P(C)}
\]
Propagation on Cliques (4)

Propagation:

\[ P(d_1) = 0.32, \text{ evidence } E = e_1, \text{ desired: } P^*(\ldots) = P(\cdot | \{e_1\}) \]

\[ P^*(c) = P(c | e_1) \]

conditional marginal distribution

\[ P^*(b, c, d) = \frac{P(b, c, d)}{P(c)} P(c | e_1) \]

multipl./division with separation prob.

\[ P(b, c, d), P^*(b, c) \]

calculate marginal distributions

\[ P^*(a, b, c) = \frac{P(a, b, c)}{P(b, c)} P(b, c | e_1) \]

multipl./division with separation prob.

\[ P^*(d_1) = P(d_1 | e_1) = 0.33 \]
Marginal distributions in the HUGIN tool.
Conditional marginal distributions with evidence $E = e_1$
Potential Representation

Let $V = \{X_j\}$ be a set of random variables $X_j : \Omega \rightarrow \text{dom}(X_j)$ and $P$ the joint distribution over $V$. Further, let

$$\{W_i \mid W_i \subseteq V, 1 \leq i \leq p\}$$

a family of subsets of $V$ with associated functions $\psi_i : \bigtimes_{X_j \in W_i} \text{dom}(X_j) \rightarrow \mathbb{R}$

It is said that $P(V)$ factorizes according $(\{W_1, \ldots, W_p\}, \{\psi_1, \ldots, \psi_p\})$ if $P(V)$ can be written as:

$$P(v) = k \cdot \prod_{i=1}^{p} \psi_i(w_i)$$

where $k \in \mathbb{R}$, $w_i$ is a realization of $W_i$ that meets the values of $v$. 
Example

\[ V = \{A, B, C\}, \quad W_1 = \{A, B\}, \quad W_2 = \{B, C\} \]

\[ \text{dom}(A) = \{a_1, a_2\} \]
\[ \text{dom}(B) = \{b_1, b_2\} \]
\[ \text{dom}(C) = \{c_1, c_2\} \]

\[ P(a, b, c) = \frac{1}{8} \]

\[ \psi_1 : \{a_1, a_2\} \times \{b_1, b_2\} \rightarrow \mathbb{R} \]
\[ \psi_2 : \{b_1, b_2\} \times \{c_1, c_2\} \rightarrow \mathbb{R} \]

\[ \psi_1(a, b) = \frac{1}{4} \]
\[ \psi_2(b, c) = \frac{1}{2} \]

\[ (\{W_1, W_2\}, \{\psi_1, \psi_2\}) \] is a potential representation of \( P \).
Factorization of a Belief Network

Let \((V, E, P)\) be an belief network and \(\{C_1, \ldots, C_p\}\) the cliques of the join tree. For every node \(v \in V\) choose a clique \(C\) such that \(v\) and all of its parents are contained in \(C\), i.e. \(\{v\} \cup c(v) \subseteq C\). The chosen clique is designated as \(f(v)\).

To arrive at a factorization \((\{C_1, \ldots, C_p\}, \{\psi_1, \ldots, \psi_p\})\) of \(P\) the factor potentials are:

\[
\psi_i(c_i) = \prod_{v : f(v) = C_i} P(v \mid c(v))
\]

Separator Sets and Residual Sets

Let \(\{C_1, \ldots, C_p\}\) be a set of cliques w.r.t. \(V\). The sets

\[
S_i = C_i \cap (C_1 \cup \cdots \cup C_{i-1}), \quad i = 1, \ldots, p, \quad S_1 = \emptyset
\]

are called separator sets with their corresponding residual sets

\[
R_i = C_i \setminus S_i
\]
Decomposition w. r. t. a Join-Tree

Given a clique ordering \( \{C_1, \ldots, C_p\} \) that satisfies the RIP, we can easily conclude the following separation statements:

\[
R_i \perp \perp (C_1 \cup \cdots \cup C_{i-1}) \setminus S_i \mid S_i \quad \text{for } i > 1
\]

Hence, we can formulate the following factorization:

\[
P(X_1, \ldots, X_n) = \prod_{i=1}^{p} P(R_i \mid S_i),
\]

which also gives us a representation in terms of conditional probabilities (as for directed graphs before).
Example

\[ S_1 = \emptyset \]
\[ R_1 = \{ A, B, C \} \]
\[ f(A) = C_1 \]
\[ S_2 = \{ B, C \} \]
\[ R_2 = \{ D \} \]
\[ f(B) = C_1 \]
\[ S_3 = \{ C \} \]
\[ R_3 = \{ E \} \]
\[ f(C) = C_1 \]
\[ f(D) = C_2 \]
\[ f(E) = C_3 \]

\[ \psi_1(C_1) = P(A, B, C \mid \emptyset) = P(A) \cdot P(C \mid A) \cdot P(B \mid A) \]
\[ \psi_2(C_2) = P(D \mid B, C) \]
\[ \psi_3(C_3) = P(E \mid C) \]

Propagation is accomplished by sending messages across the cliques in the tree. The emerging potentials are maintained by each clique.
**Main Idea**

Incorporate evidence into the clique potentials.

Since we are dealing with a tree structure, exploit the fact that a clique “separates” all its neighboring cliques (and their respective subtrees) from each other.

Apply a message passing scheme to inform neighboring cliques about evidence.

Since we do not have edge directions, we will only need one type of message.

After having updated all cliques’ potentials, we marginalize (and normalize) to get the probabilities of single attributes.
Incorporating Evidence

Every clique $C_i$ maintains a potential function $\psi_i$.

If for an attribute $E$ some evidence $e$ becomes known, we alter all potential functions of cliques containing $E$ as follows:

$$\psi_i^*(c_i) = \begin{cases} 
0, & \text{if a value in } c_i \text{ is inconsistent with } e \\
\psi_i(c_i), & \text{otherwise}
\end{cases}$$

All other potential functions are unchanged.
In general:

- Clique $C_i$ has $q$ neighboring cliques $B_1, \ldots, B_q$.
- $C_{ij}$ is the set of cliques in the subtree containing $C_i$ after dropping the link to $B_j$.
- $X_{ij}$ is the set of attributes in the cliques of $C_{ij}$.
- $V = X_{ij} \cup X_{ji}$ (complementary sets)
- $S_{ij} = S_{ji} = C_i \cap C_j$ (not shown here)
- $R_{ij} = X_{ij} \setminus S_{ij}$ (not shown here)

Here:

- Neighbors of $C_1$: $\{C_2, C_4, C_3\}$, $C_{13} = \{C_1, C_2, C_4\}$
- $X_{13} = \{A, B, C, D, E, G\}$, $S_{13} = \{C, G\}$
- $V = X_{13} \cup X_{31} = \{A, B, C, D, E, F, G, H\}$
- $R_{13} = \{A, B, D, E\}$, $R_{31} = \{F, H\}$
Task: Calculate $P(s_{ij})$:

\[
\begin{align*}
V \setminus S_{ij} &= (X_{ij} \cup X_{ji}) \setminus S_{ij} \\
&= (X_{ij} \setminus S_{ij}) \cup (X_{ji} \setminus S_{ij}) \\
&= R_{ij} \cup R_{ji}
\end{align*}
\]

\[
\begin{align*}
V \setminus S_{13} &= (X_{13} \cup X_{31}) \setminus S_{13} \\
&= R_{13} \cup R_{31}
\end{align*}
\]

\[
\begin{align*}
V \setminus \{C, G\} &= \{A, B, D, E\} \cup \{F, H\} \\
&= \{A, B, D, E, F, H\}
\end{align*}
\]

Note: $R_{ij}$ is the set of attributes that are in $C_i$’s subtree but not in $B_j$’s. Therefore, $R_{ij}$ and $R_{ji}$ are always disjoint.
Task: Calculate $P(s_{ij})$:

$$P(s_{ij}) = \sum_{v \setminus s_{ij}} \prod_{k=1}^{m} \psi_k(c_k)$$

last slide

$$= \sum_{r_{ij} \cup r_{ji}} \prod_{k=1}^{m} \psi_k(c_k)$$

sum rule

$$= \left( \sum_{r_{ij}} \prod_{c_k \in C_{ij}} \psi_k(c_k) \right) \cdot \left( \sum_{r_{ji}} \prod_{c_k \in C_{ji}} \psi_k(c_k) \right)$$

$$= M_{ij}(s_{ij}) \cdot M_{ji}(s_{ij})$$

$M_{ij}$ is the message sent from $C_i$ to neighbor $B_j$ and vice versa.
**Task:** Calculate $P(c_i)$:

$$V \setminus C_i = \left( \bigcup_{k=1}^{q} X_{ki} \right) \setminus C_i$$

$$= \bigcup_{k=1}^{q} \left( X_{ki} \setminus C_i \right)$$

$$= \bigcup_{k=1}^{q} R_{ki}$$

**Example:**

$$V \setminus C_1 = R_{21} \cup R_{41} \cup R_{31}$$

$$\{A, D, F, H\} = \{A\} \cup \{D\} \cup \{F, H\}$$
Task: Calculate $P(c_i)$:

$$P(c_i) = \sum_{v \setminus c_i} \prod_{j=1}^{m} \psi_j(c_j)$$

Marginalization Decomposition

$$= \psi_i(c_i) \sum_{v \setminus c_i} \prod_{j \neq i} \psi_j(c_j)$$

$$= \psi_i(c_i) \sum_{r_1 \cup \ldots \cup r_q} \prod_{j \neq i} \psi_j(c_j)$$

$$= \psi_i(c_i) \left( \sum_{r_{1i} \cap c_k \in C_{1i}} \prod \psi_k(c_k) \right) \ldots \left( \sum_{r_{qi} \cap c_k \in C_{qi}} \prod \psi_k(c_k) \right)$$

$$= \psi_i(c_i) \prod_{j=1}^{q} M_{ji}(s_{ij})$$
Example: $P(c_1)$:

$$P(c_1) = \psi_1(c_1)M_{21}(s_{12})M_{41}(s_{14})M_{31}(s_{13})$$

$M_{ij}(s_{ij})$ can be simplified further (without proof):

$$M_{ij}(s_{ij}) = \sum_{r_{ij}} \prod_{c_k \in C_{ij}} \psi_k(c_k) = \sum_{c_i \setminus s_{ij}} \psi_i(c_i) \prod_{k \neq j} M_{ki}(s_{ki})$$
Input:           Join tree \((C, \Psi)\) over set of variables \(V\) and evidence \(E = e\).
Output:         The a-posteriori probability \(P(x_i | e)\) for every non-evidential \(X_i\).
Initialization: Incorporate evidence \(E = e\) into potential functions.
Iterations:
1. For every clique \(C_i\) do: For every neighbor \(B_j\) of \(C_i\) do: If \(C_i\) has received all messages from the other neighbors, calculate and send \(M_{ij}(s_{ij})\) to \(B_j\).
2. Repeat step 1 until no message is calculated.
3. Calculate the joint probability distribution for every clique:
\[
P(c_i) \propto \psi_i(c_i) \prod_{j=1}^{q} M_{ji}(s_{ij})
\]
4. For every \(X \in V\) calculate the a-posteriori probability:
\[
P(x_i | e) = \sum_{c_k \setminus x_i} P(c_k)
\]
where \(C_k\) is the smallest clique containing \(X_i\).
Example: Putting it together

Goals: Find the marginal distributions and update them when evidence $H = h_1$ becomes known.

Steps:
1. Transform network into join-tree.
2. Specify factor potentials.
3. Propagate “zero” evidence to obtain the marginals before evidence is present.
4. Update factor potentials w.r.t. the evidence and do another propagation run.
Example: Step 1: Find a Join-Tree

Join-Tree creation:
Example: Step 1: Find a Join-Tree

Join-Tree creation:
1. Moralize the graph.
Example: Step 1: Find a Join-Tree

Join-Tree creation:
1. Moralize the graph.
2. Not yet triangulated.
Example: Step 1: Find a Join-Tree

Join-Tree creation:
1. Moralize the graph.
2. Triangulate the graph.
Example: Step 1: Find a Join-Tree

Join-Tree creation:
1. Moralize the graph.
2. Triangulate the graph.
3. Identify the maximal cliques.
Example: Step 1: Find a Join-Tree

Example Bayesian network

One of the join trees
Example: Step 2: Specify the Factor Potentials

Decomposition of $P(A, B, C, D, E, F, G, H)$:

$$P(a, b, c, d, e, f, g, h) = \prod_{i=1}^{5} \Psi_i(c_i) = \Psi_1(b, c, e, g) \cdot \Psi_2(a, b, c) \cdot \Psi_3(c, f, g) \cdot \Psi_4(b, d) \cdot \Psi_5(g, f, h)$$

Where to get the factor potentials from?
Example: Step 2: Specify the Factor Potentials

As long as the factor potentials multiply together as on the previous slide, we are free to choose them.

**Option 1:** A factor potential of clique $C_i$ is the product of all conditional probabilities of all node families properly contained in $C_i$:

$$\Psi_i(c_i) = 1 \cdot \prod_{\{X_i\} \cup Y_i \subseteq C_i \land \text{parents}(X_i) = Y_i} P(x_i | y_i)$$

The 1 stresses that if no node family satisfies the product condition, we assign a constant 1 to the potential.

**Option 2:** Choose potentials from the decomposition formula:

$$P(\bigcup_{i=1}^{n} C_i) = \frac{\prod_{i=1}^{n} P(C_i)}{\prod_{j=1}^{m} P(S_j)}$$
Example: Step 2: Specify the Factor Potentials

**Option 1:** Factor potentials according to the conditional distributions of the node families of the underlying Bayesian network:

\[
\begin{align*}
\Psi_1(b, c, e, g) &= P(e \mid b, c) \cdot P(g \mid e, b) \\
\Psi_2(a, b, c) &= P(b \mid a) \cdot P(c \mid a) \cdot P(a) \\
\Psi_3(c, f, g) &= P(f \mid c) \\
\Psi_4(b, d) &= P(d \mid b) \\
\Psi_5(g, f, h) &= P(h \mid g, f)
\end{align*}
\]

(This assignment of factor potentials is used in this example.)

**Option 2:** Factor potentials chosen from the join-tree decomposition:

\[
\begin{align*}
\Psi_1(b, c, e, g) &= P(b, e \mid c, g) \\
\Psi_2(a, b, c) &= P(a \mid b, c) \\
\Psi_3(c, f, g) &= P(c \mid f, g) \\
\Psi_4(b, d) &= P(d \mid b) \\
\Psi_5(g, f, h) &= P(h, g, f)
\end{align*}
\]
Encoded independence statements:
Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.
Encoded independence statements:

Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

\[ A \perp\!\!\!\!\!\!\!\perp D, E, F, G, H \mid B, C \]
Example: Closer Look on Option 2: Separation in a Join-Tree

Encoded independence statements:

Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

\[ A \perp\perp D, E, F, G, H \mid B, C \]
\[ D \perp\perp A, C, E, F, G, H \mid B \]
Example: Closer Look on Option 2: Separation in a Join-Tree

Encoded independence statements:

Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

\[
\begin{align*}
A \perp D, E, F, G, H \mid B, C \\
D \perp A, C, E, F, G, H \mid B \\
A, B, E, D \perp F, H \mid G, C
\end{align*}
\]
Encoded independence statements:
Given any separator, the variables in the cliques on one side become independent of the variables in the cliques on the other side.

\[
A \perp\!
\perp D, E, F, G, H \mid B, C \\
D \perp\!
\perp A, C, E, F, G, H \mid B \\
A, B, E, D \perp\!
\perp F, H \mid G, C \\
H \perp\!
\perp A, B, C, D, E \mid F, G
\]
The four separation statements translate into the following independence statements:

\[
\begin{align*}
A \indep D, E, F, G, H \mid B, C & \iff P(A \mid B, C, D, E, F, G, H) = P(A \mid B, C) \\
D \indep A, C, E, F, G, H \mid B & \Rightarrow P(D \mid B, C, E, F, G, H) = P(D \mid B) \\
A, B, E, D \indep F, H \mid G, C & \Rightarrow P(B, E \mid G, C, F, H) = P(B, E \mid G, C) \\
H \indep A, B, C, D, E \mid F, G & \Rightarrow P(C \mid F, G, H) = P(C \mid F, G)
\end{align*}
\]

According to the chain rule we always have the following relation:

\[
\]
The four separation statements translate into the following independence statements:

\[
\begin{align*}
A \perp \perp D, E, F, G, H \mid B, C & \Leftrightarrow P(A \mid B, C, D, E, F, G, H) = P(A \mid B, C) \\
D \perp \perp A, C, E, F, G, H \mid B & \Rightarrow P(D \mid B, C, E, F, G, H) = P(D \mid B) \\
A, B, E, D \perp \perp F, H \mid G, C & \Rightarrow P(B, E \mid G, C, F, H) = P(B, E \mid G, C) \\
H \perp \perp A, B, C, D, E \mid F, G & \Rightarrow P(C \mid F, G, H) = P(C \mid F, G)
\end{align*}
\]

Exploiting the above independencies yields:

\[
P(A, B, C, D, E, F, G, H) = P(A \mid B, C) \cdot P(D \mid B) \cdot P(B, E \mid C, G) \cdot P(C \mid F, G) \cdot P(F, G, H)
\]
The four separation statements translate into the following independence statements:

\[
\begin{align*}
A \perp \perp D, E, F, G, H \mid B, C & \iff P(A \mid B, C, D, E, F, G, H) = P(A \mid B, C) \\
D \perp \perp A, C, E, F, G, H \mid B & \Rightarrow P(D \mid B, C, E, F, G, H) = P(D \mid B) \\
A, B, E, D \perp \perp F, H \mid G, C & \Rightarrow P(B, E \mid G, C, F, H) = P(B, E \mid G, C) \\
H \perp \perp A, B, C, D, E \mid F, G & \Rightarrow P(C \mid F, G, H) = P(C \mid F, G)
\end{align*}
\]

Getting rid of the conditions results in the final decomposition equation:

\[
P(A, B, C, D, E, F, G, H) = P(A \mid B, C)P(D \mid B)P(B, E \mid C, G)P(C \mid F, G)P(F, G, H) = P(C_1)P(C_2)P(C_3)P(C_4)P(C_5) \]

\[
\frac{P(A, B, C)P(D, B)P(B, E, C, G)P(C, F, G)P(F, G, H)}{P(B, C)P(B)P(C, G)P(F, G)} \]

\[
= \frac{P(S_{12})P(S_{14})P(S_{13})P(S_{35})}{P(S_{12})P(S_{14})P(S_{13})P(S_{35})}
\]
Example: Step 3: Messages to be sent for Propagation

According to the join-tree propagation algorithm, the probability distributions of all clique instantiations $c_i$ is calculated as follows:

$$P(c_i) \propto \Psi_i(c_i) \prod_{j=1}^{q} M_{ji}(s_{ij})$$

Spelt out for our example, we get:

- $P(c_1) = P(b, c, e, g) = \Psi_1(b, c, e, g) \cdot M_{21}(b, c) \cdot M_{31}(c, g) \cdot M_{41}(b)$
- $P(c_2) = P(a, b, c) \propto \Psi_2(a, b, c) \cdot M_{12}(b, c)$
- $P(c_3) = P(c, f, g) \propto \Psi_3(c, f, g) \cdot M_{13}(c, g) \cdot M_{53}(f, g)$
- $P(c_4) = P(b, d) \propto \Psi_4(b, d) \cdot M_{14}(b)$
- $P(c_5) = P(f, g, h) \propto \Psi_5(f, g, h) \cdot M_{35}(f, g)$

The $\propto$-symbol indicates that the right-hand side may not add up to one. In that case we just normalize.
Example: Step 3: Message Computation Order

The structure of the join-tree imposes a partial ordering according to which the messages need to be computed:

\[ M_{41}(b) = \sum_d \Psi_4(b, d) \]

\[ M_{53}(f, g) = \sum_h \Psi_5(f, g, h) \]

\[ M_{21}(b, c) = \sum_a \Psi_2(a, b, c) \]

\[ M_{31}(c, g) = \sum_f \Psi_3(c, f, g) M_{53}(f, g) \]

\[ M_{13}(c, g) = \sum_{b,e} \Psi_1(b, c, e, g) M_{21}(b, c) M_{41}(b) \]

\[ M_{12}(b, c) = \sum_{e,g} \Psi_2(b, c, e, g) M_{31}(c, g) M_{41}(b) \]

\[ M_{14}(b) = \sum_{c,e,g} \Psi_1(b, c, e, g) M_{21}(b, c) M_{31}(c, g) \]

\[ M_{35}(f, g) = \sum_c \Psi_3(c, f, g) M_{13}(c, g) \]

Arrows represent is-needed-for relations. Messages on the same level can be computed in any order. Messages are computed level-wise from top to bottom.
Example: Step 3: Initialization (Potential Layouts)

<table>
<thead>
<tr>
<th>$\Psi_2$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>$b_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_2$</td>
<td>$b_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Psi_1$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$</td>
<td>$c_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$e_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$e_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$e_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_2$</td>
<td>$c_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$e_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$e_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$e_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Psi_3$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c_1$</td>
<td>$f_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$f_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>$c_2$</td>
<td>$f_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$f_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Psi_4$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td>$b_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\Psi_5$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_1$</td>
<td>$g_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$g_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$g_1$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$g_2$</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Example: Step 3: Initialization (Potential Values)
Example: Step 3: Initialization (Sending Messages)

\[ M_{21} = (b_1, b_2) = (0.06, 0.10, 0.40, 0.44) \]

\[ M_{41} = (b_1, b_2) = (1, 1) \]
Example: Step 3: Initialization (Sending Messages)

M_{21} = \begin{pmatrix} b_1, c_1 & b_1, c_2 & b_2, c_1 & b_2, c_2 \end{pmatrix} = (0.06, 0.10, 0.40, 0.44)

M_{41} = \begin{pmatrix} b_1 & b_2 \end{pmatrix} = (1, 1)

M_{13} = \begin{pmatrix} c_1, g_1 & c_1, g_2 & c_2, g_1 & c_2, g_2 \end{pmatrix} = (0.254, 0.206, 0.290, 0.250)

M_{35} = \begin{pmatrix} f_1, g_1 & f_1, g_2 & f_2, g_1 & f_2, g_2 \end{pmatrix} = (0.14, 0.12, 0.40, 0.33)
Example: Step 3: Initialization (Sending Messages)

\[ M_{21} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2) \]
\[ M_{41} = (b_1, b_2) \]
\[ M_{13} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2) \]
\[ M_{35} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2) \]
\[ M_{53} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2) \]
\[ M_{31} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2) \]
Example: Step 3: Initialization (Sending Messages)

\[ M_{21} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2) \]

\[ M_{41} = (1, 1) \]

\[ M_{13} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2) \]

\[ M_{35} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2) \]

\[ M_{53} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2) \]

\[ M_{31} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2) \]

\[ M_{12} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2) \]

\[ M_{14} = (0.16, 0.84) \]
Example: Step 3: Initialization Complete
Example: Step 4: Evidence $H = h_1$ (Altering Potentials)
Example: Step 4: Evidence $H = h_1$ (Sending Messages)

$$M_{53} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2)$$
Example: Step 4: Evidence $H = h_1$ (Sending Messages)

\[ M_{53} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2) \]
\[ M_{21} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2) \]
\[ M_{41} = (b_1, b_2) \]
Example: Step 4: Evidence $H = h_1$ (Sending Messages)

\[
\begin{align*}
M_{53} &= (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2) \\
M_{21} &= (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2) \\
M_{41} &= (b_3) \\
M_{31} &= (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2)
\end{align*}
\]
Example: Step 4: Evidence $H = h_1$ (Sending Messages)

$$M_{53} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2)$$

$$M_{21} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2)$$

$$M_{41} = (b_1, b_2)$$

$$M_{31} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2)$$

$$M_{12} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2)$$

$$M_{14} = (b_1, b_2)$$

$$M_{13} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2)$$

$$M_{35} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2)$$
Example: Step 4: Evidence $H = h_1$ Incorporated

$$M_{53} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2)$$
$$M_{21} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2)$$
$$M_{41} = (b_1, b_2)$$
$$M_{31} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2)$$
$$M_{12} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2)$$
$$M_{14} = (b_1, b_2)$$
$$M_{13} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2)$$
$$M_{35} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2)$$
Manual creation of a reasoning system based on a graphical model:

- causal model of given domain
- conditional independence graph
- decomposition of the distribution
- evidence propagation scheme

Problem: strong assumptions about the statistical effects of causal relations.
Nevertheless this approach often yields usable graphical models.
Example 1: Genotype Determination of Danish Jersey Cattle

Assumptions about parents: risk about misstatement

Genotype mother (dam)  Genotype father (sire)

Genotype child: 6 possible values

4 lysis values measured by photometer

Reliability of databases
Inheritance rules
Blood group determination
Example 1: Genotype Determination of Danish Jersey Cattle

Danish Jersey Cattle Blood Type Determination

21 attributes:
1 – dam correct?
2 – sire correct?
3 – stated dam ph.gr. 1
4 – stated dam ph.gr. 2
5 – stated sire ph.gr. 1
6 – stated sire ph.gr. 2
7 – true dam ph.gr. 1
8 – true dam ph.gr. 2
9 – true sire ph.gr. 1
10 – true sire ph.gr. 2
11 – offspring ph.gr. 1
12 – offspring ph.gr. 2
13 – offspring genotype
14 – factor 40
15 – factor 41
16 – factor 42
17 – factor 43
18 – lysis 40
19 – lysis 41
20 – lysis 42
21 – lysis 43

The grey nodes correspond to observable attributes.

This graph was specified by human domain experts, based on knowledge about (causal) dependences of the variables.
Full 21-dimensional domain has $2^6 \cdot 3^{10} \cdot 6 \cdot 8^4 = 92\,876\,046\,336$ possible states. Bayesian network requires only 306 conditional probabilities.

Example of a conditional probability table (attributes 2, 9, and 5):

<table>
<thead>
<tr>
<th>sire correct</th>
<th>true sire phenogroup 1</th>
<th>stated sire phenogroup 1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
<td>V1</td>
</tr>
<tr>
<td>yes</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>yes</td>
<td>V1</td>
<td>0</td>
</tr>
<tr>
<td>yes</td>
<td>V2</td>
<td>0</td>
</tr>
<tr>
<td>no</td>
<td>F1</td>
<td>0.58</td>
</tr>
<tr>
<td>no</td>
<td>V1</td>
<td>0.58</td>
</tr>
<tr>
<td>no</td>
<td>V2</td>
<td>0.58</td>
</tr>
</tbody>
</table>

The probabilities are acquired from human domain experts or estimated from historical data.
Example 1: Genotype Determination of Danish Jersey Cattle

moral graph
(already triangulated)

join tree
Example 1: Genotype Determination of Danish Jersey Cattle

Marginal distributions before setting evidence:
Conditional distributions given evidence in the input variables:
Example 2: Item Planning at Volkswagen

Strategy of the VW Group

<table>
<thead>
<tr>
<th>Marketing strategy</th>
<th>Vehicle specification by clients</th>
<th>Bestsellers defined by manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complexity</td>
<td>Huge number of variants</td>
<td>Small number of variants</td>
</tr>
</tbody>
</table>

Vehicle specification

<table>
<thead>
<tr>
<th>Equipment</th>
<th>fastback</th>
<th>2.8 l, 150 kW</th>
<th>Type Alpha</th>
<th>4</th>
<th>leather</th>
<th>...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>car body type</td>
<td>engine</td>
<td>radio</td>
<td>doors</td>
<td>seat cover</td>
<td>...</td>
</tr>
</tbody>
</table>
Example 2: Model “Golf”

Approx. 200 equipment groups

2 to 50 items per group

Therefore more than $2^{200}$ possible vehicle specifications

Choice of valid specifications is constrained by a rule system
(10000 technical rules, plus marketing and production rules)

Example of technical rules:

**If** Engine=$e_1$ **then** Transmission=$t_3$

**If** Engine=$e_4$ and Heating=$h_2$ **then** Generator $\in \{g_3, g_4, g_5\}$
Problem Representation

**Historical Data**
Sample of produced *vehicle specifications*
(representative choice, context-dependent, e.g. Golf)

**System of Rules**
*Rules* for the validity of item combinations
(specified for a vehicle class and a planning interval)

**Prediction & Planning**
Predicted / assigned *planning data*
(production program, demands, installation rates, capacity restrictions, …)
Complexity of the Planning Problem

Equipment table

<table>
<thead>
<tr>
<th></th>
<th>Engine</th>
<th>Transmission</th>
<th>Heating</th>
<th>Generator</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>e₁</td>
<td>t₃</td>
<td>h₁</td>
<td>g₁</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>e₂</td>
<td>t₄</td>
<td>h₃</td>
<td>g₅</td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>100000</td>
<td>e₇</td>
<td>t₁</td>
<td>h₃</td>
<td>g₂</td>
<td></td>
</tr>
</tbody>
</table>

Installation rates

<table>
<thead>
<tr>
<th>Engine</th>
<th>Transmission</th>
<th>Heating</th>
<th>Generator</th>
<th></th>
<th>Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>e₁</td>
<td>t₁</td>
<td>h₁</td>
<td>g₁</td>
<td></td>
<td>0.0000012</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>

Result is a 200-dimensional, finite probability space

\[ P(\text{Engine} = e₁, \text{Transmission} = t₃) = ? \]

\[ P(\text{Heating} = h₁ \mid \text{Generator} = g₃) = ? \]

Problem of complexity!
### Solution: Decomposition into Subspaces

\[
P(E, H, T, A) = P(A \mid E, H, T) \cdot P(T \mid E, H) \cdot P(E \mid H) \cdot P(H)
\]

Here:

\[
= P(A \mid E, H) \cdot P(T \mid E) \cdot P(E) \cdot P(H)
\]
Clique Tree of the VW Bora

Rudolf Kruse, Matthias Steinbrecher, Pascal Held

Bayesian Networks

289
Typical Planning Operation: Focusing

**Application:**

- **Compute item demand**
  Calculation of installation rates of equipment combinations

- **Simulation**
  Analyze customer requirements (e.g. of persons having ordered a navigation system for a VW Polo)

**Input:** Equipment combinations

**Operation:** Compute

- the conditional network distribution and
- the probabilities of the specified equipment combinations.
Implementation and Deployment

Project leader: Intelligent System Consulting

Client server system

Server on 6–8 machines

Quadcore platform

Terabyte hard drive

Java, Linux, Oracle

WebSphere application server

Software used daily worldwide

20 developers

5000 Bayesian networks are currently used