Evidence Propagation in Bayesian Network in Real World applications
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 sub-trees) 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.
Potential Functions

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)
- $S_{ij} = S_{ji} = C_i \cap C_j$ (Separator sets)
- $R_{ij} = X_{ij} \setminus S_{ij}$ (Residual sets)

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 $P(s_{ij})$:

In general:

$$V \setminus S_{ij} = (X_{ij} \cup X_{ji}) \setminus S_{ij}$$

$$= (X_{ij} \setminus S_{ij}) \cup (X_{ji} \setminus S_{ij})$$

$$= R_{ij} \cup R_{ji}$$

Here:

$$V \setminus S_{13} = (X_{13} \cup X_{31}) \setminus S_{13}$$

$$= R_{13} \cup R_{31}$$

$$V \setminus \{C, G\} = \{A, B, D, E\} \cup \{F, H\}$$

$$= \{A, B, D, E, F, H\}$$

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 $P(s_{ij})$:

In general:

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

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.
Propagation

Task: Calculate \( P(c_i) \):

In general:
\[
V \setminus C_i = \left( \bigcup_{k=1}^{q} X_{ki} \right) \setminus C_i
\]
\[
= \bigcup_{k=1}^{q} (X_{ki} \setminus C_i)
\]
\[
= \bigcup_{k=1}^{q} R_{ki}
\]

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 $P(c_i)$:

In general:

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

Marginalization Decomposition

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

$$= \psi_i(c_i) \sum_{r_{1i} \cup \ldots \cup r_{qi}} \prod_{i \neq j} \psi_j(c_j)$$

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

$$= \psi_i(c_i) \prod_{j=1}^{q} M_{ji}(s_{ij})$$
Propagation

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):

$$M_{ij}(s_{ij}) = \sum_{r_{ij}} \prod_{c_k \in C_{ij}} \psi_k(c_k)$$

$$= \sum_{c_i \setminus s_{ij}} \psi_i(c_i) \prod_{k \neq j} M_{ki}(s_{ki})$$
Final Algorithm

**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 \(X_i\).

The \(\propto\) sign indicates that the values \(P(c_i)\) need to be normalized if their sum is not 1.
Example 1: Clique Tree Propagation

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 1: Find a Join-Tree

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

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

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

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

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

[Diagram showing a complex network of nodes labeled A, B, C, D, E, F, G, and H.]

Cliques

One of the possible join trees

Rudolf Kruse
Bayesian Networks
Example 1: 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 1: 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 $C_i$:

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

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 1: 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.

\[ A \perp D, E, F, G, H \mid B, C \]
Example 1: Closer Look on Option 2 (Separation)

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.

\[
A \perp D, E, F, G, H \mid B, C \\
D \perp A, C, E, F, G, H \mid B
\]
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.

\[
\begin{align*}
A & \perp D, E, F, G, H \mid B, C \\
D & \perp A, C, E, F, G, H \mid B \\
A, B, E, D & \perp F, H \mid G, C
\end{align*}
\]
Example 1: Closer Look on Option 2 (Separation)

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.

\[
\begin{align*}
A \perp D, E, F, G, H & \mid B, C \\
D \perp A, C, E, F, G, H & \mid B \\
A, B, E, D & \perp F, H \mid G, C \\
H \perp A, B, C, D, E & \mid F, G
\end{align*}
\]
Example 1: Closer Look on Option 2 (Decomposition)

The four separation statements translate into the following independence statements:

\[ A \perp D, E, F, G, H \mid B, C \iff P(A \mid B, C, D, E, F, G, H) = P(A \mid B, C) \]

\[ D \perp A, C, E, F, G, H \mid B \Rightarrow P(D \mid B, C, E, F, G, H) = P(D \mid B) \]

\[ A, B, E, D \perp F, H \mid G, C \Rightarrow P(B, E \mid G, C, F, H) = P(B, E \mid G, C) \]

\[ H \perp A, B, C, D, E \mid F, G \Rightarrow P(C \mid F, G, H) = P(C \mid F, G) \]

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

\[
\]
Example 1: Closer Look on Option 2 (decomposition)

The four separation statements translate into the following independence statements:

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

Exploiting the above independencies yields:

\[
P(A, B, C, D, E, F, G, H) = P(A \mid B, C) \cdot P(D \mid B) \cdot P(B, E \mid C, G) \cdot P(C \mid F, G) \cdot P(F, G, H)
\]
Example 1: Closer Look on Option 2 (Decomposition)

The four separation statements translate into the following independence statements:

\[ A \perp D, E, F, G, H \mid B, C \quad \iff \quad P(A \mid B, C, D, E, F, G, H) = P(A \mid B, C) \]
\[ D \perp A, C, E, F, G, H \mid B \quad \Rightarrow \quad P(D \mid B, C, E, F, G, H) = P(D \mid B) \]
\[ A, B, E, D \perp F, H \mid G, C \quad \Rightarrow \quad P(B, E \mid G, C, F, H) = P(B, E \mid G, C) \]
\[ H \perp A, B, C, D, E \mid F, G \quad \Rightarrow \quad P(C \mid F, G, H) = P(C \mid F, G) \]

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

\[
P(A, B, C, D, E, F, G, H) = P(A \mid B, C)P(D \mid B)P(B, E \mid C, G)P(C \mid F, G)P(F, G, H)
\]

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

\[
= \frac{P(C_1)P(C_2)P(C_3)P(C_4)P(C_5)}{P(S_{12})P(S_{14})P(S_{13})P(S_{35})}
\]
Example 1: 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 1: 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 1: Initialization (Potential Layouts)

### Potential Layouts

**Abstract**

In this example, we illustrate the concept of potential layouts in the context of Bayesian Networks. The diagrams represent different potential models that can be used to initialize a Bayesian Network.

### Table 1: Initialization (Potential Layouts)

<table>
<thead>
<tr>
<th>$\psi_1$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1$</td>
<td></td>
</tr>
<tr>
<td>$a_2$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\psi_2$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$</td>
<td></td>
</tr>
<tr>
<td>$b_2$</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></td>
</tr>
<tr>
<td>$c_2$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\psi_4$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_1$</td>
<td></td>
</tr>
<tr>
<td>$d_2$</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\psi_5$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_1$</td>
<td></td>
</tr>
<tr>
<td>$g_2$</td>
<td></td>
</tr>
</tbody>
</table>

### Diagram

The diagrams illustrate the relationships between different variables in the potential layouts. Each node represents a variable, and the edges between nodes indicate dependencies.

**Legend**

- $A$, $B$, $C$: Variables
- $D$, $E$, $F$: Variables
- $G$, $H$, $I$: Variables

**Probabilities**

- $P(A)$, $P(B)$, $P(C)$, $P(D)$, $P(E)$, $P(F)$
- $P(G|A)$, $P(H|A)$, $P(I|A)$, $P(J|A)$

**References**

Rudolf Kruse, Bayesian Networks.
Example 1: Initialization (Potential Values)
Example 1: Initialization (Sending Messages)

\[
\begin{align*}
\psi_2 & | P \\
\hline
a_1 & \\
b_1 & c_1 0.036 \\
c_1 & 0.336 \\
b_2 & c_1 0.144 \\
c_2 & 0.012 \\
\hline
a_2 & \\
b_1 & c_1 0.228 \\
c_1 & 0.028 \\
b_2 & c_1 0.252 \\
c_2 & 0.108 \\
\hline
\end{align*}
\]

\[
\begin{align*}
\psi_4 & | P \\
\hline
b_1 & d_1 0.4 \\
d_1 & 0.6 \\
b_2 & d_1 0.7 \\
d_2 & 0.3 \\
\hline
\end{align*}
\]

\[
\begin{align*}
\psi_3 & | P \\
\hline
c_1 & \\
f_1 & g_1 0.1 \\
g_1 & 0.4 \\
g_2 & 0.3 \\
\hline
c_2 & \\
f_1 & g_1 0.4 \\
g_1 & 0.3 \\
g_2 & 0.6 \\
\hline
\end{align*}
\]

\[
\begin{align*}
\psi_5 & | P \\
\hline
f_1 & \\
h_1 & 0.2 \\
h_2 & 0.8 \\
\hline
g_1 & \\
h_1 & 0.5 \\
h_2 & 0.5 \\
\hline
f_2 & \\
h_1 & 0.4 \\
h_2 & 0.6 \\
\hline
\end{align*}
\]

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

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

\[
\begin{align*}
\psi_1 &= \begin{array}{ccc}
\text{b1} & g1 & 0.190 \\
& g2 & 0.010 \\
\text{c1} & g1 & 0.380 \\
& g2 & 0.020 \\
\text{c2} & g1 & 0.240 \\
& g2 & 0.360 \\
\text{c1} & g1 & 0.210 \\
& g2 & 0.090 \\
\text{c2} & g1 & 0.350 \\
& g2 & 0.350 \\
\text{c1} & g1 & 0.070 \\
& g2 & 0.030 \\
\text{c2} & g1 & 0.450 \\
& g2 & 0.450 \\
\end{array} \\
\psi_2 &= \begin{array}{ccc}
\text{b1} & c1 & 0.036 \\
& c2 & 0.084 \\
\text{c1} & c1 & 0.144 \\
& c2 & 0.336 \\
\text{b1} & c1 & 0.028 \\
& c2 & 0.012 \\
\text{b2} & c1 & 0.252 \\
& c2 & 0.108 \\
\end{array} \\
\psi_3 &= \begin{array}{ccc}
\text{c1} & g1 & 0.1 \\
& g2 & 0.1 \\
\text{f1} & g1 & 0.4 \\
& g2 & 0.6 \\
\text{f2} & g1 & 0.7 \\
& g2 & 0.3 \\
\end{array} \\
\psi_4 &= \begin{array}{ccc}
\text{b1} & d1 & 0.4 \\
& d2 & 0.6 \\
\text{b2} & d1 & 0.7 \\
& d2 & 0.3 \\
\end{array} \\
\psi_5 &= \begin{array}{ccc}
\text{g1} & h1 & 0.2 \\
& h2 & 0.8 \\
\text{g2} & h1 & 0.5 \\
& h2 & 0.5 \\
\text{g1} & h1 & 0.6 \\
& h2 & 0.7 \\
\end{array} \\
\end{align*}
\]

M_{21} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2) = (0.06, 0.10, 0.40, 0.44)

M_{41} = (b_1, b_2) = (1, 1)

M_{13} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2) = (0.254, 0.206, 0.290, 0.250)

M_{35} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2) = (0.14, 0.12, 0.40, 0.33)
Example 1: Initialization (Sending Messages)

$$\psi_1$$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c_1$</td>
<td>$e_1$</td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td>$e_2$</td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
<tr>
<td>$c_2$</td>
<td>$e_1$</td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td>$e_2$</td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
</tbody>
</table>

$$\psi_2$$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_1$</td>
<td>$b_1$</td>
<td>c1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
<td>c1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
<tr>
<td>$a_2$</td>
<td>$b_1$</td>
<td>c1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
<tr>
<td></td>
<td>$b_2$</td>
<td>c1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
</tbody>
</table>

$$\psi_3$$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_3$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$c_1$</td>
<td>$f_1$</td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
<tr>
<td></td>
<td>$f_2$</td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
<tr>
<td>$c_2$</td>
<td>$f_1$</td>
<td>g1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>g2</td>
</tr>
</tbody>
</table>

$$\psi_4$$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_4$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_1$</td>
<td>$d_1$</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>$d_2$</td>
<td>0.6</td>
</tr>
<tr>
<td>$b_2$</td>
<td>$d_1$</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>$d_2$</td>
<td>0.3</td>
</tr>
</tbody>
</table>

$$\psi_5$$

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\psi_5$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$f_1$</td>
<td>$g_1$</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>$h_1$</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>$h_2$</td>
<td>0.5</td>
</tr>
<tr>
<td>$f_2$</td>
<td>$g_1$</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>$h_1$</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>$h_2$</td>
<td>0.3</td>
</tr>
</tbody>
</table>

$$M_{21} = \begin{pmatrix} b_1, c_1 \ b_1, c_2 \ b_2, c_1 \ b_2, c_2 \end{pmatrix}$$

$$M_{41} = \begin{pmatrix} b_1 \ b_2 \end{pmatrix}$$

$$M_{13} = \begin{pmatrix} c_1, g_1 \ c_1, g_2 \ c_2, g_1 \ c_2, g_2 \end{pmatrix}$$

$$M_{35} = \begin{pmatrix} f_1, g_1 \ f_1, g_2 \ f_2, g_1 \ f_2, g_2 \end{pmatrix}$$

$$M_{53} = \begin{pmatrix} f_1, g_1 \ f_1, g_2 \ f_2, g_1 \ f_2, g_2 \end{pmatrix}$$

$$M_{31} = \begin{pmatrix} c_1, g_1 \ c_1, g_2 \ c_2, g_1 \ c_2, g_2 \end{pmatrix}$$
Example 1: Initialization (Sending Messages)

\[ M_{21} = (b_1, c_1, b_2, c_1, b_2, c_2, b_3, c_2, b_4, c_4) \]
\[ M_{41} = (b_1, b_2) \]
\[ M_{13} = (c_1, g_1, c_1, g_2, c_2, g_1, c_2, g_2) \]
\[ M_{43} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2) \]
\[ M_{53} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2) \]
\[ M_{12} = (b_1, c_1, b_2, c_2, b_3, c_3, b_4, c_4) \]
\[ M_{14} = (b_1, b_2) \]
Example 1: Initialization Complete

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
$\Psi_2$ & $P$ & \\
\hline
$\Psi_2$ & $P$ & \\
\hline
$\Psi_4$ & $P$ & \\
\hline
$\Psi_3$ & $P$ & \\
\hline
$\Psi_5$ & $P$ & \\
\hline
\end{tabular}
\end{table}

\begin{align*}
M_{21} &= (b_1, c_1, 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) \\
M_{12} &= (b_1, c_1, b_2, c_1, b_2, c_2) \\
M_{14} &= (b_1, b_2) \\
\end{align*}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|}
\hline
&P & A & B & C & D & E & F & G & H \\
\hline
1 & 0.6000 & 0.1600 & 0.4600 & 0.6520 & 0.2144 & 0.2620 & 0.5448 & 0.4842 \\
2 & 0.4000 & 0.8400 & 0.4500 & 0.3480 & 0.7856 & 0.7380 & 0.4552 & 0.5158 \\
\hline
\end{tabular}
\end{table}
Example 1: Evidence $H = h_1$ (Altering Potentials)
Example 1: Evidence $H = h_1$ (Sending Messages)

$$M_{53} = \begin{pmatrix} f_{1,g1} & f_{1,g2} & f_{2,g1} & f_{2,g2} \\ 0.2 & 0.5 & 0.4 & 0.7 \end{pmatrix}$$
### Example 1: Step 4: Evidence $H = h_1$ (Sending Messages)

#### Bayesian Network Diagram

- **Node $C_2$:** $ABC$
- **Node $C_4$:** $BD$
- **Node $C_3$:** $EG$
- **Node $C_5$:** $GH$

#### Evidence Table

<table>
<thead>
<tr>
<th>$\psi_1$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_1$</td>
<td>$c_1$ 0.190</td>
</tr>
<tr>
<td></td>
<td>$c_2$ 0.010</td>
</tr>
<tr>
<td>$b_2$</td>
<td>$c_1$ 0.320</td>
</tr>
<tr>
<td></td>
<td>$c_2$ 0.480</td>
</tr>
<tr>
<td>$c_1$</td>
<td>$e_1$ 0.380</td>
</tr>
<tr>
<td></td>
<td>$e_2$ 0.020</td>
</tr>
<tr>
<td>$c_2$</td>
<td>$e_1$ 0.240</td>
</tr>
<tr>
<td></td>
<td>$e_2$ 0.360</td>
</tr>
</tbody>
</table>

#### Node $C_1$
- $b_1$ $d_1$ 0.4
- $b_2$ $d_1$ 0.7
- $b_2$ $d_2$ 0.3

#### Node $C_5$
- $f_1$ $g_1$ 0.1
- $f_2$ $g_1$ 0.4

#### Node $C_3$
- $b_1$ $c_1$ 0.036
- $b_2$ $c_1$ 0.144
- $b_1$ $c_2$ 0.336
- $b_2$ $c_2$ 0.012
- $b_1$ $e_1$ 0.268
- $b_2$ $e_1$ 0.252

#### Node $C_4$
- $f_1$ $h_1$ 0.2
- $f_2$ $h_1$ 0.4

#### Node $C_5$
- $f_1$ $g_1$ 0.320
- $f_2$ $g_1$ 0.240
- $f_1$ $g_2$ 0.480
- $f_2$ $g_2$ 0.360

#### Node $C_1$
- $b_1$ $d_1$ 0.4
- $b_2$ $d_1$ 0.7
- $b_2$ $d_2$ 0.3

#### Node $C_3$
- $f_1$ $g_1$ 0.1
- $f_2$ $g_1$ 0.4

#### Node $C_5$
- $f_1$ $h_1$ 0.4
- $f_2$ $h_1$ 0.7
- $f_1$ $h_2$ 0
- $f_2$ $h_2$ 0

#### Evaluation:

- $M_{53} = (f_1, g_1, f_1, g_2, f_2, g_1, f_2, g_2, 0.2, 0.3, 0.4, 0.7)$
- $M_{21} = (b_1, c_1, b_1, c_2, b_2, c_1, b_2, c_2)$
- $M_{41} = (1, 1)$
Example 1: Evidence $H = h_1$ (Sending Messages)

\[
\begin{array}{c|c|c}
\psi_2 & a_1 & a_2 \\
\hline
b_1 & c_1 & 0.036 \\
& c_2 & 0.084 \\
& b_1 & 0.28 \\
& b_2 & 0.12 \\
& b_2 & c_1 & 0.252 \\
& c_2 & 0.108 \\
\end{array}
\]

\[
\begin{array}{c|c|c}
\psi_3 & f_1 & c_1 \\
\hline
& g_1 & 0.1 \\
& g_2 & 0.1 \\
& c_2 & 0.4 \\
& g_2 & 0.4 \\
\end{array}
\]

\[
\begin{array}{c|c|c}
\psi_4 & b_1 & b_2 \\
\hline
& d_1 & 0.4 \\
& d_2 & 0.6 \\
& d_1 & 0.7 \\
& d_2 & 0.3 \\
\end{array}
\]

\[
\begin{array}{c|c|c}
\psi_5 & f_1 & f_2 \\
\hline
& g_1 & h_1 \ 0.2 \\
& & h_2 \ 0 \\
& g_2 & h_1 \ 0.5 \\
& & h_2 \ 0 \\
& g_1 & h_1 \ 0.4 \\
& & h_2 \ 0 \\
& g_2 & h_1 \ 0.7 \\
& & h_2 \ 0 \\
\end{array}
\]

\[
M_{53} = \begin{pmatrix} f_1, g_1 & f_1, g_2 & f_2, g_1 & f_2, g_2 \end{pmatrix} = \begin{pmatrix} 0.2 & 0.5 & 0.4 & 0.7 \end{pmatrix}
\]

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

\[
M_{41} = \begin{pmatrix} b_1 & b_2 \end{pmatrix}
\]

\[
M_{31} = \begin{pmatrix} c_1, g_1 & c_1, g_2 & c_2, g_1 & c_2, g_2 \end{pmatrix} = \begin{pmatrix} 0.38 & 0.68 & 0.32 & 0.62 \end{pmatrix}
\]
Example 1: Evidence $H = h_1$ (Sending Messages)

\[
\begin{align*}
M_{53} &= \left( f_{1.1}, f_{1.2}, f_{2.1} \right) \\
M_{21} &= \left( b_{1.1}, b_{1.2}, b_{2.1} \right) \\
M_{41} &= \left( b_{1}, b_{2} \right) \\
M_{31} &= \left( c_{1.1}, c_{1.2}, c_{2.1}, c_{2.2}, c_{2.2} \right) \\
M_{12} &= \left( b_{1.1}, b_{1.2}, b_{2.1}, b_{2.2} \right) \\
M_{14} &= \left( b_{1}, b_{2} \right) \\
M_{13} &= \left( c_{1.1}, c_{1.2}, c_{2.1}, c_{2.2}, c_{2.2} \right) \\
M_{35} &= \left( f_{1.1}, f_{1.2}, f_{2.1} \right)
\end{align*}
\]
Example 1: Evidence $H = h_1$ Incorporated

$$M_{53} = \left( f_{1.11} = f_{1.12} = f_{2.11} = f_{2.12} \right)$$

$$M_{21} = \left( b_{1.1}, b_{1.2}, b_{2.1}, b_{2.2} \right)$$

$$M_{41} = \left( b_{1.1}, b_{2.2} \right)$$

$$M_{31} = \left( c_{1.11}, c_{1.12}, c_{2.11}, c_{2.12}, c_{2.21}, c_{2.22} \right)$$

$$M_{12} = \left( b_{1.1}, b_{1.2}, b_{2.1}, b_{2.2}, b_{2.1}, b_{2.2} \right)$$

$$M_{14} = \left( b_{1.1}, b_{2.2} \right)$$

$$M_{13} = \left( c_{1.11}, c_{1.12}, c_{2.11}, c_{2.12}, c_{2.21}, c_{2.22} \right)$$

$$M_{35} = \left( f_{1.11} = f_{1.12}, f_{2.11} = f_{2.12} \right)$$

<table>
<thead>
<tr>
<th>$P$</th>
<th>$A$</th>
<th>$B$</th>
<th>$C$</th>
<th>$D$</th>
<th>$E$</th>
<th>$F$</th>
<th>$G$</th>
<th>$H$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\cdot$1</td>
<td>0.5888</td>
<td>0.1557</td>
<td>0.4884</td>
<td>0.6533</td>
<td>0.1899</td>
<td>0.1828</td>
<td>0.3916</td>
<td>1.0000</td>
</tr>
<tr>
<td>$\cdot$2</td>
<td>0.4112</td>
<td>0.8443</td>
<td>0.5116</td>
<td>0.3467</td>
<td>0.8101</td>
<td>0.8172</td>
<td>0.6084</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
HUGIN's Solution
There are several exact inference methods for Bayesian Networks beside the clique tree propagation such as variable elimination or recursive conditioning. These algorithms have all complexity that is exponential with networks tree width. Exact inference is NP-hard.

In very large applications it is often useful to introduce topological structural constraints or restrictions on conditional probabilities, i.e. bounded variance algorithms.

There are also several approximate inference methods.
Example 2: 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

See the Paper „Blood group determination of Danish Jersey Cattle in the F-blood group system“ by Lene Kolind Rasmussen for the details.
Example 2: 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 dependences between variables.
Example 2: Genotype Determination of Danish Jersey Cattle

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 phenogroup 1</th>
<th>F1</th>
<th>V1</th>
<th>V2</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>F1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>yes</td>
<td>V1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>yes</td>
<td>V2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>no</td>
<td>F1</td>
<td>0.58</td>
<td>0.10</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>V1</td>
<td>0.58</td>
<td>0.10</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>V2</td>
<td>0.58</td>
<td>0.10</td>
<td>0.32</td>
<td></td>
</tr>
</tbody>
</table>

The probabilities are acquired from human domain experts or estimated from historical data.
Example 2: Genotype Determination of Danish Jersey Cattle

moral graph
(already triangulated)

join tree
Example 2: Genotype Determination of Danish Jersey Cattle

Marginal distributions before setting evidence:

Bayesian Networks
Conditional distributions given evidence in the input variables:
**Example 3: Property planning - Volkswagen**

<table>
<thead>
<tr>
<th>Property family</th>
<th>Car body</th>
<th>Motor</th>
<th>Radio</th>
<th>Doors</th>
<th>Seat cover</th>
<th>Makeup mirror</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hatch-back</td>
<td>2.8 L 150 kW Otto</td>
<td>Type alpha</td>
<td>4</td>
<td>Leather, Type L3</td>
<td>yes</td>
<td></td>
</tr>
</tbody>
</table>

**Complexity**
- About 200 variables
- Typically 4 to 8, but up to 150 possible instances per variable
- More than $2^{200}$ possible combinations available

Example 3: Handling the System of Technical Rules

- 10000 Technical Rules for Item Combinations, e.g.
  
  \[
  \text{IF } \text{Motor} = m_4 \ \text{AND} \ \text{Heating} = h_1
  \]

  \[
  \text{THEN} \ \text{Generator} \in \{g_3, g_4, g_5\}
  \]

- Technical Rules can be seen as Constraints, e.g. 3-dimensional relations

- The Rules are often 6-dimensional, sometimes more than 10 dimensions

- 500000 marketing oriented rules
Example 3: Property planning

- Goals:
  Model possible (part-relevant) property combinations
  Support demand forecasts for all (part-relevant) property combinations

- Planning intervals: short-term, mid-term

- Context: Model groups, Planning intervals

- Daily: 5000 planning scenarios handled by 350 planners worldwide

Assistant System for Handling the Planner‘s Knowledge about the Installation Rates of Property Combinations
Calculation of part demands

Compute the installation rate of a given item combination

Simulation

Analyze customers' preferences with respect to those persons who use a navigation system in a VW Polo

Marketing and Sales stipulation

Installation rate of Navigation system increase from 20% to 30%

Capacity Restrictions

Maximum availability of seat coverings in leather is 5000

In the language of the philosopher Gärdenfors: An agent (planner) is in a Belief State, he is using the belief change operations Contraction (Focusing) and Revision
Example 3: Qualitative and Quantitative Information about Property Planning

**Data**
- Vehicle orders
- Specifications of built vehicles (sample)

**Knowledge**
- Installation rules
- Combinability of properties

**Context**
- Model group
- Planning horizon

**Forecast/Planning**
forecast/set plan data
(frequencies, required quantities, capacities (restrictions), production plans, open purchase order quantities ...)

Diagram:
- Data
- Knowledge
- Context
- Forecast/Planning

?
Example 3: Requirements of Users

Planners (2008)
- Explicit, Sound, Transparent Model
- Explanation of the Results
- Answers to the Questions in real time (seconds)
- Automatic Integration of New Information into the Model

Law (since 2018)
EU’s General Data Protection Regulation (GDPR) includes a right to explanation: “The use of AI tools should be transparent, explainable, fair, and empirically sound while fostering accountability.”
Example 3: Markov Network for VW Bora

186 variables, 174 cliques
Example 3: System EPL (EigenschaftsPPlanung) at VW

Project leader: Intelligent System Consulting (PD Dr. habil. Jörg Gebhardt)

In worldwide use: 15 developers, 350 planners

Different planning responsibilities with individual workflow
- Assessment of demand for approx. 40 planning intervals (weeks, months)
- 5000 different Markov networks in use daily
Important Topics for Real Applications

**Fusion of Qualitative and Quantitative Knowledge**
Data, Rule Systems, Conditional Independence Statements, Contexts, etc.

**Learning Models from Data**
- Parameters (e.g. Conditional Probabilities) and Structure (e.g. DAG, Cliques)
- Model Change in the light of new Information (rules, probabilities)
- Handling Inconsistencies and Missing Values, Modelling Causalities
- Scalability, Transparency, Audability, Accountability, Accuracy,...

**Decision Making**
Decisions under Uncertainty, Uncertainty Quantification (Epistemic vs Aleatoric),

**Trustworthy Solutions of AI Solutions**
- **Ethical, Lawful, Robust (from a technical perspective and in its social environment)**
- Safety, Fairness, Non-discrimination, Privacy and Data Governance,
- Human Agency and Oversight, Societal and Environmental Well-Being...