Belief network inference

Three main approaches to determine posterior distributions in belief networks:

➤ Exploiting the structure of the network to eliminate (sum out) the non-observed, non-query variables one at a time.

➤ Search-based approaches that enumerate some of the possible worlds, and estimate posterior probabilities from the worlds generated.

➤ Stochastic simulation where random cases are generated according to the probability distributions.
Suppose $B$ is Boolean ($B = true$ is $b$ and $B = false$ is $\neg b$)

$$P(C|A)$$

$$= P(C \land b|A) + P(C \land \neg b|A)$$

$$= P(C|b \land A)P(b|A) + P(C|\neg b \land A)P(\neg b|A)$$

$$= P(C|b)P(b|A) + P(C|\neg b)P(\neg b|A)$$

$$= \sum_B P(C|B)P(B|A)$$

We can compute the probability of some of the variables by summing out the other variables.
Factors

A factor is a representation of a function from a tuple of random variables into a number.

We will write factor $f$ on variables $X_1, \ldots, X_j$ as $f(X_1, \ldots, X_j)$.

We can assign some or all of the variables of a factor:

- $f(X_1 = v_1, X_2, \ldots, X_j)$, where $v_1 \in \text{dom}(X_1)$, is a factor on $X_2, \ldots, X_j$.

- $f(X_1 = v_1, X_2 = v_2, \ldots, X_j = v_j)$ is a number that is the value of $f$ when each $X_i$ has value $v_i$.

The former is also written as $f(X_1, X_2, \ldots, X_j)_{X_1 = v_1}$, etc.
**Example factors**

![Table of example factors](image)

\[
\begin{array}{ccc|c}
X & Y & Z & \text{val} \\
\hline
\text{t} & \text{t} & \text{t} & 0.1 \\
\text{t} & \text{t} & \text{f} & 0.9 \\
\text{t} & \text{f} & \text{t} & 0.2 \\
\text{f} & \text{t} & \text{t} & 0.4 \\
\text{f} & \text{t} & \text{f} & 0.6 \\
\text{f} & \text{f} & \text{t} & 0.3 \\
\text{f} & \text{f} & \text{f} & 0.7 \\
\end{array}
\]

\[
\begin{array}{cc|c}
Y & Z & \text{val} \\
\hline
\text{t} & \text{t} & 0.1 \\
\text{t} & \text{f} & 0.9 \\
\text{f} & \text{t} & 0.2 \\
\text{f} & \text{f} & 0.8 \\
\end{array}
\]

\[
r(X=t, Y, Z) = 0.8
\]

\[
r(X=t, Y, Z=f) = 0.8
\]

\[
r(X=t, Y=f, Z=f) = 0.8
\]

© David Poole, Alan Mackworth, Randy Goebel, and Oxford University Press 1998-2002
Multiplying factors

The **product** of factor $f_1(\bar{X}, \bar{Y})$ and $f_2(\bar{Y}, \bar{Z})$, where $\bar{Y}$ are the variables in common, is the factor $(f_1 \times f_2)(\bar{X}, \bar{Y}, \bar{Z})$ defined by:

$$(f_1 \times f_2)(\bar{X}, \bar{Y}, \bar{Z}) = f_1(\bar{X}, \bar{Y})f_2(\bar{Y}, \bar{Z}).$$
# Multiplying factors example

### $f_1$: $AB$

<table>
<thead>
<tr>
<th>$A$</th>
<th>$B$</th>
<th>val</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td>0.1</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>0.9</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>0.2</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>0.8</td>
</tr>
</tbody>
</table>

### $f_2$: $BC$

<table>
<thead>
<tr>
<th>$B$</th>
<th>$C$</th>
<th>val</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td>0.3</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>0.7</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>0.6</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>0.4</td>
</tr>
</tbody>
</table>

### $f_1 \times f_2$: $ABC$

<table>
<thead>
<tr>
<th>$A$</th>
<th>$B$</th>
<th>$C$</th>
<th>val</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td>t</td>
<td>0.03</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>f</td>
<td>0.07</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>t</td>
<td>0.54</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>t</td>
<td>0.36</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>f</td>
<td>0.06</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>t</td>
<td>0.14</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>f</td>
<td>0.48</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>f</td>
<td>0.32</td>
</tr>
</tbody>
</table>

© David Poole, Alan Mackworth, Randy Goebel, and Oxford University Press 1998-2002
We can sum out a variable, say $X_1$ with domain $\{v_1, \ldots, v_k\}$, from factor $f(X_1, \ldots, X_j)$, resulting in a factor on $X_2, \ldots, X_j$ defined by:

$$
(\sum_{X_1} f)(X_2, \ldots, X_j) = f(X_1 = v_1, \ldots, X_j) + \cdots + f(X_1 = v_k, \ldots, X_j)
$$
### Multiplying factors example

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>val</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>0.03</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>f</td>
<td>f</td>
<td>0.07</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>t</td>
<td>t</td>
<td>0.54</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>t</td>
<td>t</td>
<td>0.06</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>f</td>
<td>f</td>
<td>0.14</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>t</td>
<td>t</td>
<td>0.48</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>f</td>
<td>f</td>
<td>0.32</td>
</tr>
</tbody>
</table>

\[ f_3 : \]

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>C</th>
<th>val</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>t</td>
<td>t</td>
<td>0.57</td>
</tr>
<tr>
<td>t</td>
<td>f</td>
<td>f</td>
<td>0.43</td>
</tr>
<tr>
<td>f</td>
<td>t</td>
<td>t</td>
<td>0.54</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>f</td>
<td>0.46</td>
</tr>
</tbody>
</table>

\[ \sum_B f_3 : \]

© David Poole, Alan Mackworth, Randy Goebel, and Oxford University Press 1998-2002
If we want to compute the posterior probability of $Z$ given evidence $Y_1 = v_1 \land \ldots \land Y_j = v_j$:

$$P(Z|Y_1 = v_1, \ldots, Y_j = v_j) = \frac{P(Z, Y_1 = v_1, \ldots, Y_j = v_j)}{P(Y_1 = v_1, \ldots, Y_j = v_j)} = \frac{P(Z, Y_1 = v_1, \ldots, Y_j = v_j)}{\sum_Z P(Z, Y_1 = v_1, \ldots, Y_j = v_j)}.$$ 

So the computation reduces to the probability of $P(Z, Y_1 = v_1, \ldots, Y_j = v_j)$. 

We normalize at the end.
Probability of a conjunction

Suppose the variables of the belief network are $X_1, \ldots, X_n$.

To compute $P(Z, Y_1 = v_1, \ldots, Y_j = v_j)$, we sum out the other variables, $Z_1, \ldots, Z_k = \{X_1, \ldots, X_n\} - \{Z\} - \{Y_1, \ldots, Y_j\}$.

We order the $Z_i$ into an elimination ordering.

\[
P(Z, Y_1 = v_1, \ldots, Y_j = v_j)
= \sum_{Z_k} \cdots \sum_{Z_1} P(X_1, \ldots, X_n)_{Y_1 = v_1, \ldots, Y_j = v_j}.
\]

\[
= \sum_{Z_k} \cdots \sum_{Z_1} \prod_{i=1}^{n} P(X_i | \pi_{X_i})_{Y_1 = v_1, \ldots, Y_j = v_j}.
\]
Computing sums of products

Computation in belief networks reduces to computing the sums of products.

➤ How can we compute $ab + ac$ efficiently?
Computing sums of products

Computation in belief networks reduces to computing the sums of products.

➤ How can we compute $ab + ac$ efficiently?

➤ Distribute out the $a$ giving $a(b + c)$
Computing sums of products

Computation in belief networks reduces to computing the sums of products.

➤ How can we compute $ab + ac$ efficiently?

➤ Distribute out the $a$ giving $a(b + c)$

➤ How can we compute $\sum_{Z_1} \prod_{i=1}^{n} P(X_i | \pi_{X_i})$ efficiently?
Computing sums of products

Computation in belief networks reduces to computing the sums of products.

➤ How can we compute $ab + ac$ efficiently?

➤ Distribute out the $a$ giving $a(b + c)$

➤ How can we compute $\sum_{Z_1} \prod_{i=1}^{n} P(X_i|\pi_{X_i})$ efficiently?

➤ Distribute out those factors that don’t involve $Z_1$. 

© David Poole, Alan Mackworth, Randy Goebel, and Oxford University Press 1998-2002
Variable elimination algorithm

To compute $P(Z|Y_1 = v_1 \land \ldots \land Y_j = v_j)$:

➤ Construct a factor for each conditional probability.

➤ Set the observed variables to their observed values.

➤ Sum out each of the other variables (the $\{Z_1, \ldots, Z_k\}$) according to some elimination ordering.

➤ Multiply the remaining factors. Normalize by dividing the resulting factor $f(Z)$ by $\sum_Z f(Z)$. 

© David Poole, Alan Mackworth, Randy Goebel, and Oxford University Press 1998-2002
Summing out a variable

To sum out a variable $Z_j$ from a product $f_1, \ldots, f_k$ of factors:

- Partition the factors into
  - those that don’t contain $Z_j$, say $f_1, \ldots, f_i$,
  - those that contain $Z_j$, say $f_{i+1}, \ldots, f_k$

We know:

$$
\sum_{Z_j} f_1 \times \cdots \times f_k = f_1 \times \cdots \times f_i \times \left( \sum_{Z_j} f_{i+1} \times \cdots \times f_k \right).
$$

- Explicitly construct a representation of the rightmost factor. Replace the factors $f_{i+1}, \ldots, f_k$ by the new factor.
Variable elimination example

\[ P(A) \] \[ \rightarrow \] \[ f_1(B) \] 
\[ P(B|A) \]
\[ P(C) \]
\[ P(D|BC) \]
\[ P(E|C) \]
\[ P(F|D) \]
\[ P(G|FE) \]
\[ P(H|G) \] \[ \rightarrow \] \[ f_3(G) \] 
\[ P(I|G) \] \[ \rightarrow \] \[ f_4(G) \] 

© David Poole, Alan Mackworth, Randy Goebel, and Oxford University Press 1998-2002