Bayesian belief networks: Inference

Milos Hauskrecht
milos@cs.pitt.edu
5329 Sennott Square

Bayesian belief networks (general)

Two components: $B = (S, \Theta_S)$
- **Directed acyclic graph**
  - Nodes correspond to random variables
  - (Missing) links encode independences
- **Parameters**
  - Local conditional probability distributions for every variable-parent configuration
  
  $$P(X_i \mid pa(X_i))$$

  Where:
  $pa(X_i)$ - stand for parents of $X_i$

\[
\begin{array}{ccc}
B & E & T & F \\
T T & 0.95 & 0.05 \\
T F & 0.94 & 0.06 \\
F T & 0.29 & 0.71 \\
F F & 0.001 & 0.999 \\
\end{array}
\]
**Full joint distribution in BBNs**

**Full joint distribution** is defined in terms of local conditional distributions (obtained via the chain rule):

\[
P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid pa(X_i))
\]

\[
P(B, E, A, J, M) = P(J \mid A)P(M \mid A)P(A \mid B, E)P(B)P(E)
\]

**Parameter complexity problem**

- In the BBN the **full joint distribution** is defined as:
  \[
P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid pa(X_i))
\]
- **What did we save?**
  
  **Alarm example:** 5 binary (True, False) variables

  - **# of parameters of the full joint:**
    \[2^5 = 32\]
    One parameter is for free:
    \[2^5 - 1 = 31\]
  - **# of parameters of the BBN:**
    \[2^3 + 2(2^2) + 2(2) = 20\]
    One parameter in every conditional is for free:
    \[2^2 + 2(2) + 2(1) = 10\]
Inference in Bayesian network

• **Bad news:**
  – Exact inference problem in BBNs is NP-hard (Cooper)
  – Approximate inference is NP-hard (Dagum, Luby)

• **But** very often we can achieve significant improvements

• Assume our Alarm network

![Alarm network diagram]

• Assume we want to compute: \( P(J = T) \)

Inference in Bayesian networks

**Computing:** \( P(J = T) \)

**Approach 1. Blind approach.**

• Sum out all un-instantiated variables from the full joint,
• express the joint distribution as a product of conditionals

\[
P(J = T) = \sum_{b \in \{T,F\}} \sum_{e \in \{T,F\}} \sum_{a \in \{T,F\}} \sum_{m \in \{T,F\}} P(B = b, E = e, A = a, J = T, M = m)
\]

\[
= \sum_{b \in \{T,F\}} \sum_{e \in \{T,F\}} \sum_{a \in \{T,F\}} \sum_{m \in \{T,F\}} P(J = T \mid A = a) P(M = m \mid A = a) P(A = a \mid B = b, E = e) P(B = b) P(E = e)
\]

**Computational cost:**
Number of additions: 15
Number of products: 16*4=64
Inference in Bayesian networks

Approach 2. Interleave sums and products

- Combines sums and product in a smart way (multiplications by constants can be taken out of the sum)

\[ P(J = T) = \]
\[ = \sum_{bct,F} \sum_{ect,F} \sum_{act,F} \sum_{wct,F} P(J = T \mid A = a)P(M = m \mid A = a)P(A = a \mid B = b, E = e)P(B = b)P(E = e) \]
\[ = \sum_{bct,F} \sum_{ect,F} \sum_{act,F} P(J = T \mid A = a)P(M = m \mid A = a)P(B = b) \left( \sum_{ecl,F} P(A = a \mid B = b, E = e)P(E = e) \right) \]
\[ = \sum_{act,F} P(J = T \mid A = a) \left( \sum_{mct,F} P(M = m \mid A = a) \right) \left( \sum_{bct,F} P(B = b) \right) \left( \sum_{ecl,F} P(A = a \mid B = b, E = e)P(E = e) \right) \]

**Computational cost:**

- Number of additions: 1+2*[1+1+2*1]=9
- Number of products: 2*[2+2*(1+2*1)]=16

---

Inference in Bayesian network

- **Exact inference algorithms:**
  - Variable elimination
  - Recursive decomposition (Cooper, Darwiche)
  - Symbolic inference (D’Ambrosio)
  - Belief propagation algorithm (Pearl)
  - Clustering and joint tree approach (Lauritzen, Spiegelhalter)
  - Arc reversal (Olmsted, Schachter)

- **Approximate inference algorithms:**
  - Monte Carlo methods:
    - Forward sampling, Likelihood sampling
    - Variational methods
Variable elimination

- **Variable elimination:**
  - Similar idea but interleave sum and products one variable at the time during inference
  - E.g. Query $P(J = T)$ requires to eliminate A,B,E,M and this can be done in different order

$$P(J = T) = \sum_{b \in T,F} \sum_{e \in T,F} \sum_{a \in T,F} \sum_{m} \sum_{b} \sum_{e} P(J = T \mid A = a) P(M = m \mid A = a) P(A = a \mid B = b, E = e) P(B = b) P(E = e)$$

---

**Variable elimination**

**Assume order:** M, E, B, A to calculate $P(J = T)$

$$= \sum_{b \in T,F} \sum_{e \in T,F} \sum_{a \in T,F} \sum_{m} \sum_{b} \sum_{e} P(J = T \mid A = a) P(M = m \mid A = a) P(A = a \mid B = b, E = e) P(B = b) P(E = e)$$
Variable elimination

Assume order: M, E, B, A to calculate $P(J = T)$

$$
= \sum_{b_{cl,T}} \sum_{e_{cl,T}} \sum_{a_{cl,T}} \sum_{m_{cl,T}} P(J = T \mid A = a) P(M = m \mid A = a) P(A = a \mid B = b, E = e) P(B = b) P(E = e)
$$

$$
= \sum_{b_{cl,T}} \sum_{e_{cl,T}} \sum_{a_{cl,T}} P(J = T \mid A = a) P(A = a \mid B = b, E = e) P(B = b) P(E = e) \left[ \sum_{m_{cl,T}} P(M = m \mid A = a) \right]
$$

M. Hauskrecht
Variable elimination

Assume order: M, E, B, A to calculate \( P(J = T) \)

\[
\tau_1(A = a, B = b) = \sum_{m \in F} P(A = m | A = a) P(B = b | A = m) P(E = e) + \sum_{m \in F} P(A = m | A = a) P(B = b | A = m) P(E = e)
\]

\[
\tau_1(A = a, B = b) = \sum_{m \in F} P(A = m | A = a) P(B = b | A = m) P(E = e) + \sum_{m \in F} P(A = m | A = a) P(B = b | A = m) P(E = e)
\]

M. Hauskrecht
Variable elimination

Assume order: M, E, B, A to calculate \( P(J = T) \)

\[
= \sum_{b \in T} \sum_{e \in T} \sum_{a \in T} \sum_{m \in T} P(J = T | A = a) P(M = m | A = a) P(A = a | B = b, E = e) P(B = b) P(E = e)
\]

\[
= \sum_{b \in T} \sum_{e \in T} \sum_{a \in T} P(J = T | A = a) P(A = a | B = b, E = e) P(B = b) P(E = e) \left[ \sum_{m \in T} P(M = m | A = a) \right]
\]

\[
= \sum_{a \in T} \sum_{b \in T} \sum_{e \in T} P(J = T | A = a) P(A = a | B = b, E = e) P(B = b) P(E = e) \left[ \sum_{m \in T} P(M = m | A = a) \right]
\]

\[
= \sum_{a \in T} \sum_{b \in T} P(J = T | A = a) P(B = b) \tau_i(A = a, B = b)
\]

---

Variable elimination

Assume order: M, E, B, A to calculate \( P(J = T) \)

\[
= \sum_{b \in T} \sum_{e \in T} \sum_{a \in T} \sum_{m \in T} P(J = T | A = a) P(M = m | A = a) P(A = a | B = b, E = e) P(B = b) P(E = e)
\]

\[
= \sum_{b \in T} \sum_{e \in T} \sum_{a \in T} P(J = T | A = a) P(A = a | B = b, E = e) P(B = b) P(E = e) \left[ \sum_{m \in T} P(M = m | A = a) \right]
\]

\[
= \sum_{a \in T} \sum_{b \in T} \sum_{e \in T} P(J = T | A = a) P(A = a | B = b, E = e) P(B = b) P(E = e) \left[ \sum_{m \in T} P(M = m | A = a) \right]
\]

\[
= \sum_{a \in T} \sum_{b \in T} P(J = T | A = a) P(B = b) \tau_i(A = a, B = b)
\]

\[
= \sum_{a \in T} P(J = T | A = a) \left[ \sum_{b \in T} P(B = b) \tau_i(A = a, B = b) \right]
\]

---

M. Hauskrecht
Variable elimination

Assume order: M, E, B, A to calculate \( P(J = T) \)

\[
= \sum_{bcF} \sum_{cF} \sum_{aF} P(J = T \mid A = a)P(M = m \mid A = a)P(A = a \mid B = b, E = e)P(B = b)P(E = e)
\]

\[
= \sum_{bcF} \sum_{cF} \sum_{aF} P(J = T \mid A = a)P(A = a \mid B = b, E = e)P(B = b)P(E = e) \left[ \sum_{mF} P(M = m \mid A = a) \right]
\]

\[
= \sum_{bcF} \sum_{cF} \sum_{aF} P(J = T \mid A = a)P(A = a \mid B = b, E = e)P(B = b)P(E = e)
\]

\[
= \sum_{bcF} \sum_{cF} \sum_{aF} P(J = T \mid A = a)P(B = b) \left[ \sum_{cF} P(A = a \mid B = b, E = e)P(E = e) \right]
\]

\[
= \sum_{bcF} \sum_{cF} \sum_{aF} P(J = T \mid A = a)P(B = b) \tau(A = a, B = b)
\]

\[
= \sum_{aF} \sum_{bcF} P(J = T \mid A = a) \left[ \sum_{cF} P(B = b) \tau(A = a, B = b) \right]
\]

\[
= \sum_{aF} \sum_{bcF} P(J = T \mid A = a) \tau(A = a)
\]

M. Hauskrecht
Variable elimination

Assume order: M, E, B, A to calculate $P(J = T)$

$$
= \sum_{bc1,F} \sum_{ecl,F} \sum_{as1,F} \sum_{mc1,F} P(J = T | A = a) P(M = m | A = a) P(A = a | B = b, E = e) P(B = b) P(E = e) \\
= \sum_{bc1,F} \sum_{ecl,F} \sum_{as1,F} P(J = T | A = a) P(A = a | B = b, E = e) P(B = b) P(E = e) \left[ \sum_{mc1,F} P(M = m | A = a) \right] \\
= \sum_{as1,F} \sum_{bc1,F} P(J = T | A = a) P(B = b) \left[ \sum_{ecl,F} P(A = a | B = b, E = e) P(E = e) \right] \\
= \sum_{as1,F} \sum_{bc1,F} P(J = T | A = a) P(B = b) \tau(A = a, B = b) \\
= \sum_{as1,F} P(J = T | A = a) \left[ \sum_{ecl,F} P(B = b) \tau(A = a, B = b) \right] \\
= \sum_{as1,F} P(J = T | A = a) \tau(A = a) = P(J = T)
$$

Inference in Bayesian network

- **Exact inference algorithms:**
  - Variable elimination
  - Recursive decomposition (Cooper, Darwiche)
  - Symbolic inference (D’Ambrosio)
  - Belief propagation algorithm (Pearl)
  - Clustering and joint tree approach (Lauritzen, Spiegelhalter)
  - Arc reversal (Olmsted, Schachter)

- **Approximate inference algorithms:**
  - Monte Carlo methods:
    - Forward sampling, Likelihood sampling
    - Variational methods
Monte Carlo approaches

• MC approximation:
  – The probability is approximated using sample frequencies
  – Example:
    \[ \tilde{P}(B = T, J = T) = \frac{N_{B=T,J=T}}{N} \]
    (num samples with \(B = T, J = T\) / total # samples)

• BBN sampling:

  B
  \(\rightarrow\)
  E
  \(\rightarrow\)
  A
  \(\rightarrow\)
  J
  \(\rightarrow\)
  M

  Generate sample in a top down manner, following the links

• One sample gives one assignment of values to all variables

BBN sampling example

\begin{align*}
P(B) & \begin{array}{cc}
T & F \\
0.001 & 0.999 \\
\end{array} \\
\hline
P(E) & \begin{array}{cc}
T & F \\
0.002 & 0.998 \\
\end{array} \\
\hline
\end{align*}

\begin{align*}
P(A|B,E) & \begin{array}{|c|c|}
B & E \\
\hline
T & T \\
0.95 & 0.05 \\
T & F \\
0.94 & 0.06 \\
F & T \\
0.29 & 0.71 \\
F & F \\
0.001 & 0.999 \\
\end{array} \\
\hline
\end{align*}

\begin{align*}
P(J|A) & \begin{array}{|c|c|}
A & T \\
\hline
T & 0.90 \\
F & 0.05 \\
0.1 & 0.95 \\
\end{array} \\
\hline
\end{align*}

\begin{align*}
P(M|A) & \begin{array}{|c|c|}
A & T \\
\hline
T & 0.7 \\
F & 0.01 \\
0.3 & 0.99 \\
\end{array} \\
\hline
\end{align*}
BBN sampling example

<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.001</td>
<td>0.999</td>
</tr>
<tr>
<td>E</td>
<td>0.002</td>
<td>0.998</td>
</tr>
</tbody>
</table>

\[
P(B) \\
T | 0.001 | 0.999 \\
F | 0.002 | 0.998 \\
\]

\[
P(E) \\
T | 0.001 | 0.999 \\
F | 0.002 | 0.998 \\
\]

\[
P(A|B,E) \\
<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>T T</td>
<td>0.95</td>
<td>0.05</td>
</tr>
<tr>
<td>T F</td>
<td>0.94</td>
<td>0.06</td>
</tr>
<tr>
<td>F T</td>
<td>0.29</td>
<td>0.71</td>
</tr>
<tr>
<td>F F</td>
<td>0.001</td>
<td>0.999</td>
</tr>
</tbody>
</table>
\]

\[
P(A) \\
<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.90</td>
<td>0.1</td>
</tr>
<tr>
<td>F</td>
<td>0.05</td>
<td>0.95</td>
</tr>
</tbody>
</table>
\]

\[
P(J|M|A) \\
<table>
<thead>
<tr>
<th></th>
<th>T</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>F</td>
<td>0.01</td>
<td>0.99</td>
</tr>
</tbody>
</table>
\]
BBN sampling example

- **Burglary**
  - \( P(B) = \begin{array}{c|c} \text{T} & 0.001 \\ \text{F} & 0.999 \end{array} \)
- **Earthquake**
  - \( P(E) = \begin{array}{c|c} \text{T} & 0.002 \\ \text{F} & 0.998 \end{array} \)

- **Alarm**
  - \( P(A|B,E) = \begin{array}{c|cc} \text{B} & \text{E} & \text{T} & \text{F} \\ \text{T} & 0.95 & 0.05 \\ \text{T} & 0.94 & 0.06 \\ \text{F} & 0.29 & 0.71 \\ \text{F} & 0.001 & 0.999 \end{array} \)

- **JohnCalls**
  - \( P(J|A) = \begin{array}{c|c|c} \text{A} & \text{T} & \text{F} \\ \text{T} & 0.90 & 0.1 \\ \text{F} & 0.05 & 0.95 \end{array} \)

- **MaryCalls**
  - \( P(M|A) = \begin{array}{c|c|c} \text{A} & \text{T} & \text{F} \\ \text{T} & 0.7 & 0.3 \\ \text{F} & 0.01 & 0.99 \end{array} \)
Monte Carlo approaches

- **MC approximation of conditional probabilities:**
  - The probability is approximated using sample frequencies
  - **Example:**
    \[ \tilde{P}(B = T \mid J = T, M = F) = \frac{N_{B=T,J=T,M=F}}{N_{J=T,M=F}} \]
    # samples with $B = T, J = T, M = F$

- **Rejection sampling**
  - Generate samples from the full joint by sampling BBN
  - Use only samples that agree with the condition, the remaining samples are rejected
- **Problem:** many samples can be rejected
Likelihood weighting

**Idea:** generate only samples consistent with an evidence (or conditioning event)
  - Benefit: Avoids inefficiencies of rejection sampling

**Problem:**
- the distribution generated by enforcing the conditioning variables to set values is biased
- simple counts are not sufficient to estimate the probabilities

**Solution:**
- With every sample keep a weight with which it should count towards the estimate

\[
\bar{P}(B = T \mid J = T, M = F) = \frac{\sum_{\text{samples with } B = T, M = F \text{ and } J = T} W_{B=T,J=T,M=F}}{\sum_{\text{samples with any values of } B \text{ and } J = T, M = F} W_{B=T,J=T,M=F}}
\]

---

**BBN likelihood weighting example**

<table>
<thead>
<tr>
<th></th>
<th>P(B)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td><strong>Burglary</strong></td>
<td>0.001</td>
<td>0.999</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>P(E)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T</td>
<td>F</td>
<td></td>
</tr>
<tr>
<td><strong>Earthquake</strong></td>
<td>0.002</td>
<td>0.998</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B</th>
<th>E</th>
<th>T</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>0.95</td>
<td>0.05</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
<td>0.94</td>
<td>0.06</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>0.29</td>
<td>0.71</td>
</tr>
<tr>
<td>F</td>
<td>F</td>
<td>0.001</td>
<td>0.999</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A</th>
<th>T</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.90</td>
<td>0.1</td>
</tr>
<tr>
<td>F</td>
<td>0.05</td>
<td>0.95</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>A</th>
<th>T</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>F</td>
<td>0.01</td>
<td>0.99</td>
</tr>
</tbody>
</table>

J = T (set !!!)
M = F (set !!!)
BBN likelihood weighting example

<table>
<thead>
<tr>
<th></th>
<th>( P(B) )</th>
<th>( P(E) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>0.001</td>
<td>0.002</td>
</tr>
<tr>
<td>( F )</td>
<td>0.999</td>
<td>0.998</td>
</tr>
</tbody>
</table>

\[
P(A|B,E)\]

<table>
<thead>
<tr>
<th>( B )</th>
<th>( E )</th>
<th>( T )</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>( T )</td>
<td>0.95</td>
<td>0.05</td>
</tr>
<tr>
<td>( T )</td>
<td>( F )</td>
<td>0.94</td>
<td>0.06</td>
</tr>
<tr>
<td>( F )</td>
<td>( T )</td>
<td>0.29</td>
<td>0.71</td>
</tr>
<tr>
<td>( F )</td>
<td>( F )</td>
<td>0.001</td>
<td>0.999</td>
</tr>
</tbody>
</table>

\[
P(J|A)\]

<table>
<thead>
<tr>
<th>( A )</th>
<th>( T )</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>0.90</td>
<td>0.1</td>
</tr>
<tr>
<td>( F )</td>
<td>0.05</td>
<td>0.95</td>
</tr>
</tbody>
</table>

\[
P(M|A)\]

<table>
<thead>
<tr>
<th>( A )</th>
<th>( T )</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>( F )</td>
<td>0.01</td>
<td>0.99</td>
</tr>
</tbody>
</table>

\( J = T \) (set !!!)
\( M = F \) (set !!!)
BBN likelihood weighting example

<table>
<thead>
<tr>
<th>Burglary</th>
<th>Earthquake</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(B)</td>
<td>P(E)</td>
</tr>
<tr>
<td>T F</td>
<td>T F</td>
</tr>
<tr>
<td>0.001</td>
<td>0.999</td>
</tr>
<tr>
<td>0.002</td>
<td>0.998</td>
</tr>
</tbody>
</table>

| Alarm    | P(A|B,E)   |
|----------|-----------|
| B E T F  |           |
| T T 0.95 | T F 0.94  |
| T F 0.29 | T F 0.001 |
| F T 0.71 | F F 0.999 |

<table>
<thead>
<tr>
<th>JohnCalls</th>
<th>MaryCalls</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(J</td>
<td>A)</td>
</tr>
<tr>
<td>A T F</td>
<td>A T F</td>
</tr>
<tr>
<td>T 0.90</td>
<td>T 0.7</td>
</tr>
<tr>
<td>F 0.05</td>
<td>F 0.01</td>
</tr>
<tr>
<td>0.95</td>
<td>0.3</td>
</tr>
<tr>
<td>0.05</td>
<td>0.999</td>
</tr>
</tbody>
</table>

J = T (set !!!)
M = F (set !!!)
BBN likelihood weighting example

**Evidence J=T, M=F in combination with B=T, E=F, A=T**
weight = 0.9*0.3 = 0.27
BBN likelihood weighting example

Second sample

\[
\begin{array}{c|cc}
\text{Burglary} & \text{T} & \text{F} \\
\hline
\text{P}(\text{B}) & 0.001 & 0.999 \\
\end{array}
\]

\[
\begin{array}{c|cc}
\text{Earthquake} & \text{T} & \text{F} \\
\hline
\text{P}(\text{E}) & 0.002 & 0.998 \\
\end{array}
\]

\[
\begin{array}{c|cc|cc}
\text{Alarm} & \text{B} & \text{E} & \text{T} & \text{F} \\
\hline
\text{B} & \text{E} & \text{T} & \text{T} & 0.95 & 0.05 \\
\text{T} & \text{F} & 0.94 & 0.06 \\
\text{F} & \text{T} & 0.29 & 0.71 \\
\text{F} & \text{F} & 0.001 & 0.999 \\
\end{array}
\]

\[
\begin{array}{c|cc|cc}
\text{JohnCalls} & \text{A} & \text{T} & \text{F} \\
\hline
\text{A} & \text{T} & 0.90 & 0.1 \\
\text{F} & 0.05 & 0.95 \\
\end{array}
\]

\[
\begin{array}{c|cc|cc}
\text{MaryCalls} & \text{A} & \text{T} & \text{F} \\
\hline
\text{A} & \text{T} & 0.7 & 0.3 \\
\text{F} & 0.01 & 0.99 \\
\end{array}
\]

J = T (set !!!)  M = F (set !!!)
BBN likelihood weighting example

Second sample

<table>
<thead>
<tr>
<th>Burglary</th>
<th>Earthquake</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(B)</td>
<td>P(E)</td>
</tr>
<tr>
<td>T</td>
<td>F</td>
</tr>
<tr>
<td>0.001</td>
<td>0.999</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Alarm</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>T</td>
</tr>
<tr>
<td>T</td>
</tr>
<tr>
<td>F</td>
</tr>
<tr>
<td>F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>JohnCalls</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(J</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>T</td>
</tr>
<tr>
<td>F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MaryCalls</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(M</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>T</td>
</tr>
<tr>
<td>F</td>
</tr>
</tbody>
</table>

\[ J = T \text{ (set !!!)} \]
\[ M = F \text{ (set !!!)} \]
BBN likelihood weighting example

Second sample

Evidence \( J=T, M=F \)
in combination with \( B=F, E=F, A=F \)

\[
\text{weight} = 0.05 \times 0.99 = 0.0495
\]

Likelihood weighting

- Assume we have generated the following M samples:

  \[
  \begin{array}{ccc
  F & F & F \\
  F & F & T \\
  T & F & F \\
  F & F & F \\
  T & F & F \end{array} \]

How to make the samples consistent? Weight each sample by probability with which it agrees with the conditioning evidence \( P(e) \).

\[
\tilde{P}(B = T \mid J = T, M = F) = \frac{\sum_{\text{samples with } B = T, M = F \text{ and } J = T} W_{B = T, J = T, M = F}}{\sum_{\text{samples with any value of } B \text{ and } J = T, M = F} W_{B = T, J = T, M = F}}
\]