Advanced Machine Learning- Exercise 3

Graphical Models and Exact Inference

M.Sc. Ömer Sali & M.Sc. Jonathon Luiten
<lastname>@vision.rwth-aachen.de

RWTH Aachen University - Computer Vision Group
http://www.vision.rwth-aachen.de/

2019-06-05

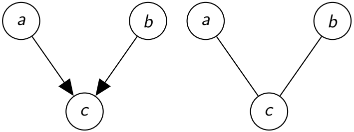
Content

1. Bayesian Networks
2. Bayes Ball, D-Separation
3. Bayes Net Toolbox
4. Factor Graphs, Sum-Product Algorithm

Content

Bayes Networks - Recap

- Two basic kinds of graphical models
  - Directed graphical models or Bayesian Networks
  - Undirected graphical models or Markov Random Fields
- Key components
  - Nodes : Random variables
  - Edges : Directed or undirected
  - The value of a random variable may be known or unknown.



Directed graph

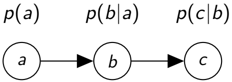
Undirected graph

Known variables

Unknown variables

Bayes Networks - Recap

- Chains of nodes:



  - Knowledge about a is expressed by the prior probability:

$$p(a)$$

  - Dependencies are expressed through conditional probabilities:

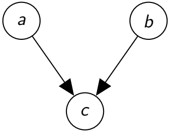
$$p(b|a), p(c|b)$$

  - Joint distribution of all three variables:

$$\begin{aligned} p(a, b, c) &= p(c|a, b)p(a, b) \\ &= p(c|b)p(b|a)p(a) \end{aligned}$$

Bayes Networks - Recap

- Convergent connections:



  - Here the value of c depends on both variables a and b.
  - This is modeled with the conditional probability:

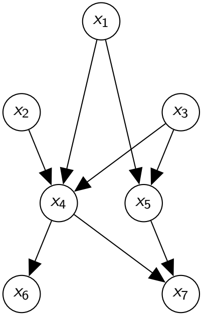
$$p(c|a, b)$$

  - Therefore, the joint probability of all three variables is given as:

$$\begin{aligned} p(a, b, c) &= p(c|a, b)p(a, b) \\ &= p(c|a, b)p(a)p(b) \end{aligned}$$

Bayes Networks - Recap

- Computing the joint probability



$$
\begin{aligned}
p(x_1, \ldots, x_7) = {} & p(x_1)p(x_2)p(x_3) \\
& p(x_4|x_1, x_2, x_3) \\
& p(x_5|x_1, x_3) \\
& p(x_6|x_4)p(x_7|x_4, x_5)
\end{aligned}
$$

General factorization

$$
p(\mathbf{x}) = \prod_{k=1}^{K} p(x_k|pa_k)
$$

We can directly read off the factorization of the joint from the network structure!

Bayes Networks - Question

- Elections are coming up!
- Political Parties: Party A, Party B, Party C
- Promised spending: low(l), moderate(m), high(h)

- Party A favors health care ($h = 0.8$, $m = 0.15$).
- Party A equally likely to promise spending in all 3 categories in education.

- Party B favors education ($h = 0.7$, $m = 0.15$).
- Party B equally likely to promise spending in all 3 categories in health care.

- Party C makes lots of promises ($h = 0.8$, $l = 0.2$) in both areas

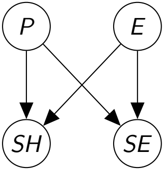
Bayes Networks - Solution

- P: political party.
- SE: promised spending in education.
- SH: promised spending in health care.
- E: election is imminent.

Bayes Network - Solution (cont.)

P	p(P)
A	0.33
B	0.33
B	0.33

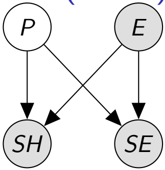
E	p(E)
T	0.5
F	0.5



P	E	P(SH = h|P, E)	P(SH = m|P, E)	P(SH = l|P, E)
A	T	0.8	0.15	0.05
B	T	0.33	0.33	0.33
C	T	0.8	0.0	0.2
A	F	0.01	0.1	0.89
B	F	0.01	0.1	0.89
C	F	0.01	0.1	0.89

P	E	P(SE = h|P, E)	P(SE = m|P, E)	P(SE = l|P, E)
A	T	0.33	0.33	0.33
B	T	0.7	0.15	0.15
C	T	0.8	0.0	0.2
A	F	0.01	0.1	0.89
B	F	0.01	0.1	0.89
C	F	0.01	0.1	0.89

Bayes Networks - Solution (cont.)

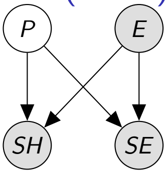


$$p(P=?|SH=m,SE=h,E=T)=\frac{\text{p(P=?, SH=m, SE=h, E=T)}}{\text{p(SH=m, SE=h, E=T)}}$$

$$\text{p(P=?, SH=m, SE=h, E=T)} = p(P=?)p(E=T)\,p(SE=h|P=?,E=T)\,p(SH=m|P=?,E=T)$$

$$\text{p(SH=m,SE=h,E=T)} = p(E=T)\sum_{\mathbf{p}\in P}\Big[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})\,p(SE=h|E=T,P=\mathbf{p})\Big]$$

Bayes Networks - Solution (cont.)



$$p(P = ?|SH = m, SE = h, E = T) = \frac{\text{p(P=?, SH=m, SE=h, E=T)}}{\text{p(SH=m, SE=h, E=T)}}$$

$$\text{p(P=?, SH=m, SE=h, E=T)} = p(P = ?)\cancel{p(E = T)}$$
$$p(SE = h|P = ?, E = T)$$
$$p(SH = m|P = ?, E = T)$$

$$\text{p(SH=m,SE=h,E=T)} = \cancel{p(E = T)}$$
$$\sum_{\mathbf{p} \in P} \Big[p(P = \mathbf{p}) p(SH = m|E = T, P = \mathbf{p})$$
$$p(SE = h|E = T, P = \mathbf{p}) \Big]$$

Bayes Networks - Solution

$$P(P = ?| \dots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p} \in P} \Big[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p}) \Big]}$$

Bayes Networks - Solution

$$P(P = ?|\ldots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p} \in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$P(P = A|\ldots) = \frac{p(P = A)p(SE = h|P = A, E = T)p(SH = m|P = A, E = T)}{\sum_{\mathbf{p} \in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

Bayes Networks - Solution

$$P(P = ?| \dots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p} \in P} \left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p}) \right]}$$

$$P(P = A| \dots) = \frac{p(P = A)p(SE = h|P = A, E = T)p(SH = m|P = A, E = T)}{\sum_{\mathbf{p} \in P} \left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p}) \right]}$$

$$= \frac{0.33 \cdot 0.33 \cdot 0.15}{0.33 \cdot 0.15 \cdot 0.33 + 0.33 \cdot 0.33 \cdot 0.7 + 0.33 \cdot 0 \cdot 0.8}$$

Bayes Networks - Solution

$$P(P = ?|\ldots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p} \in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$P(P = A|\ldots) = \frac{p(P = A)p(SE = h|P = A, E = T)p(SH = m|P = A, E = T)}{\sum_{\mathbf{p} \in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.15}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}}$$

Bayes Networks - Solution

$$P(P = ?|\dots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$P(P = A|\dots) = \frac{p(P = A)p(SE = h|P = A, E = T)p(SH = m|P = A, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.15}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}} = \mathbf{0.17647}$$

Bayes Networks - Solution

$$P(P = ?|\ldots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$P(P = A|\ldots) = \frac{p(P = A)p(SE = h|P = A, E = T)p(SH = m|P = A, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.15}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}} = \mathbf{0.17647}$$

$$P(P = B|\ldots) = \frac{p(P = B)p(SE = h|P = B, E = T)p(SH = m|P = B, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

Bayes Networks - Solution

$$P(P = ?|\dots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$P(P = A|\dots) = \frac{p(P = A)p(SE = h|P = A, E = T)p(SH = m|P = A, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.15}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}} = \mathbf{0.17647}$$

$$P(P = B|\dots) = \frac{p(P = B)p(SE = h|P = B, E = T)p(SH = m|P = B, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$= \frac{0.33 \cdot 0.33 \cdot 0.7}{0.33 \cdot 0.15 \cdot 0.33 + 0.33 \cdot 0.33 \cdot 0.7 + 0.33 \cdot 0 \cdot 0.8}$$

Bayes Networks - Solution

$$P(P=?|\dots) = \frac{p(P=?)p(SE=h|P=?,E=T)p(SH=m|P=?,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$

$$P(P=A|\dots) = \frac{p(P=A)p(SE=h|P=A,E=T)p(SH=m|P=A,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$

$$= \frac{\cancel{0.33}\cdot\cancel{0.33}\cdot 0.15}{\cancel{0.33}\cdot 0.15\cdot\cancel{0.33}+\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7+\cancel{0.33\cdot 0\cdot 0.8}} = \mathbf{0.17647}$$

$$P(P=B|\dots) = \frac{p(P=B)p(SE=h|P=B,E=T)p(SH=m|P=B,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$

$$= \frac{\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7}{\cancel{0.33}\cdot 0.15\cdot\cancel{0.33}+\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7+\cancel{0.33\cdot 0\cdot 0.8}}$$

Bayes Networks - Solution

$$P(P = ?|\ldots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p} \in P} \left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p}) \right]}$$

$$P(P = A|\ldots) = \frac{p(P = A)p(SE = h|P = A, E = T)p(SH = m|P = A, E = T)}{\sum_{\mathbf{p} \in P} \left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p}) \right]}$$
$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.15}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}} = \mathbf{0.17647}$$

$$P(P = B|\ldots) = \frac{p(P = B)p(SE = h|P = B, E = T)p(SH = m|P = B, E = T)}{\sum_{\mathbf{p} \in P} \left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p}) \right]}$$
$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.7}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}} = \mathbf{0.8235}$$

Bayes Networks - Solution

$$P(P = ?|\dots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p}\in P}\Big[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\Big]}$$

$$P(P = A|\dots) = \frac{p(P = A)p(SE = h|P = A, E = T)p(SH = m|P = A, E = T)}{\sum_{\mathbf{p}\in P}\Big[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\Big]}$$
$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.15}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}} = \mathbf{0.17647}$$

$$P(P = B|\dots) = \frac{p(P = B)p(SE = h|P = B, E = T)p(SH = m|P = B, E = T)}{\sum_{\mathbf{p}\in P}\Big[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\Big]}$$
$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.7}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}} = \mathbf{0.8235}$$

$$P(P = C|\dots) = \frac{p(P = C)p(SE = h|P = C, E = T)p(SH = m|P = C, E = T)}{\sum_{\mathbf{p}\in P}\Big[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\Big]}$$

Bayes Networks - Solution

$$P(P=?|\ldots) = \frac{p(P=?)p(SE=h|P=?,E=T)p(SH=m|P=?,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$

$$P(P=A|\ldots) = \frac{p(P=A)p(SE=h|P=A,E=T)p(SH=m|P=A,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$

$$= \frac{\cancel{0.33}\cdot\cancel{0.33}\cdot 0.15}{\cancel{0.33}\cdot 0.15\cdot\cancel{0.33}+\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7+\cancel{0.33\cdot 0\cdot 0.8}} = \mathbf{0.17647}$$

$$P(P=B|\ldots) = \frac{p(P=B)p(SE=h|P=B,E=T)p(SH=m|P=B,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$

$$= \frac{\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7}{\cancel{0.33}\cdot 0.15\cdot\cancel{0.33}+\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7+\cancel{0.33\cdot 0\cdot 0.8}} = \mathbf{0.8235}$$

$$P(P=C|\ldots) = \frac{p(P=C)p(SE=h|P=C,E=T)p(SH=m|P=C,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$

$$= \frac{0.33\cdot 0\cdot 0.8}{0.33\cdot 0.15\cdot 0.33+0.33\cdot 0.33\cdot 0.7+0.33\cdot 0\cdot 0.8}$$

Bayes Networks - Solution

$$P(P = ?|\ldots) = \frac{p(P = ?)p(SE = h|P = ?, E = T)p(SH = m|P = ?, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$P(P = A|\ldots) = \frac{p(P = A)p(SE = h|P = A, E = T)p(SH = m|P = A, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.15}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}} = \mathbf{0.17647}$$

$$P(P = B|\ldots) = \frac{p(P = B)p(SE = h|P = B, E = T)p(SH = m|P = B, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$= \frac{\cancel{0.33} \cdot \cancel{0.33} \cdot 0.7}{\cancel{0.33} \cdot 0.15 \cdot \cancel{0.33} + \cancel{0.33} \cdot \cancel{0.33} \cdot 0.7 + \cancel{0.33 \cdot 0 \cdot 0.8}} = \mathbf{0.8235}$$

$$P(P = C|\ldots) = \frac{p(P = C)p(SE = h|P = C, E = T)p(SH = m|P = C, E = T)}{\sum_{\mathbf{p}\in P}\left[p(P = \mathbf{p})p(SH = m|E = T, P = \mathbf{p})p(SE = h|E = T, P = \mathbf{p})\right]}$$

$$= \frac{\cancel{0.33 \cdot 0 \cdot 0.8}}{0.33 \cdot 0.15 \cdot 0.33 + 0.33 \cdot 0.33 \cdot 0.7 + 0.33 \cdot 0 \cdot 0.8}$$

Bayes Networks - Solution

$$P(P=?|\dots) = \frac{p(P=?)p(SE=h|P=?,E=T)p(SH=m|P=?,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$

$$P(P=A|\dots) = \frac{p(P=A)p(SE=h|P=A,E=T)p(SH=m|P=A,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$
$$= \frac{\cancel{0.33}\cdot\cancel{0.33}\cdot 0.15}{\cancel{0.33}\cdot 0.15\cdot\cancel{0.33} + \cancel{0.33}\cdot\cancel{0.33}\cdot 0.7 + \cancel{0.33\cdot 0\cdot 0.8}} = \mathbf{0.17647}$$

$$P(P=B|\dots) = \frac{p(P=B)p(SE=h|P=B,E=T)p(SH=m|P=B,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$
$$= \frac{\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7}{\cancel{0.33}\cdot 0.15\cdot\cancel{0.33} + \cancel{0.33}\cdot\cancel{0.33}\cdot 0.7 + \cancel{0.33\cdot 0\cdot 0.8}} = \mathbf{0.8235}$$

$$P(P=C|\dots) = \frac{p(P=C)p(SE=h|P=C,E=T)p(SH=m|P=C,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$
$$= \frac{\cancel{0.33\cdot 0\cdot 0.8}}{0.33\cdot 0.15\cdot 0.33 + 0.33\cdot 0.33\cdot 0.7 + 0.33\cdot 0\cdot 0.8} = \mathbf{0}$$

Bayes Networks - Solution

$$P(P=?|\ldots) = \frac{p(P=?)p(SE=h|P=?,E=T)p(SH=m|P=?,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$

$$P(P=A|\ldots) = \frac{p(P=A)p(SE=h|P=A,E=T)p(SH=m|P=A,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$
$$= \frac{\cancel{0.33}\cdot\cancel{0.33}\cdot 0.15}{\cancel{0.33}\cdot 0.15\cdot\cancel{0.33}+\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7+\cancel{0.33\cdot 0\cdot 0.8}} = \mathbf{0.17647}$$

$$P(P=B|\ldots) = \frac{p(P=B)p(SE=h|P=B,E=T)p(SH=m|P=B,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$
$$= \frac{\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7}{\cancel{0.33}\cdot 0.15\cdot\cancel{0.33}+\cancel{0.33}\cdot\cancel{0.33}\cdot 0.7+\cancel{0.33\cdot 0\cdot 0.8}} = \mathbf{0.8235}$$

$$P(P=C|\ldots) = \frac{p(P=C)p(SE=h|P=C,E=T)p(SH=m|P=C,E=T)}{\sum_{\mathbf{p}\in P}\left[p(P=\mathbf{p})p(SH=m|E=T,P=\mathbf{p})p(SE=h|E=T,P=\mathbf{p})\right]}$$
$$= \frac{\cancel{0.33\cdot 0\cdot 0.8}}{0.33\cdot 0.15\cdot 0.33+0.33\cdot 0.33\cdot 0.7+0.33\cdot 0\cdot 0.8} = \mathbf{0}$$

It is highly probable that politician belongs to Party B!

Bayes Networks - Solution

- If we know that Party C is not fielding any candidates in our constituency then $p(P = A) = p(P = B) = 0.5$.

$$P(P = \mathbf{A}|\ldots) = \frac{0.5 \cdot 0.33 \cdot 0.15}{0.5 \cdot 0.15 \cdot 0.33 + 0.5 \cdot 0.33 \cdot 0.7 + 0 \cdot 0 \cdot 0.8} = 0.17647$$

$$P(P = \mathbf{B}|\ldots) = \frac{0.5 \cdot 0.33 \cdot 0.7}{0.5 \cdot 0.15 \cdot 0.33 + 0.5 \cdot 0.33 \cdot 0.7 + 0 \cdot 0 \cdot 0.8} = 0.8235$$

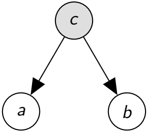
$$P(P = \mathbf{C}|\ldots) = \frac{0 \cdot 0 \cdot 0.8}{0.5 \cdot 0.15 \cdot 0.33 + 0.5 \cdot 0.33 \cdot 0.7 + 0 \cdot 0 \cdot 0.8} = 0$$

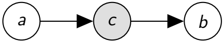
The distribution does not change!

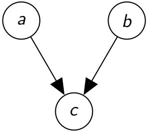
Content

D-Separation -Reminder

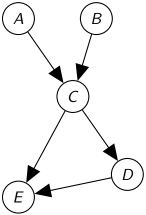
- Definition
  - Let X, Y, and Z be non-intersecting subsets of nodes in a directed graph.
  - A path from X to Y is blocked if it contains a node such that either
    - The arrows on the path meet either head-to-tail or tail-to-tail at the node, and the node is in the set Z of observed variables, or
    - The arrows meet head-to-head at the node, and neither the node, nor any of its descendants, are in the set Z.
  - If all paths from X to Y are blocked, X is said to be d-separated from Y given Z.







"Bayes Ball"- Recap

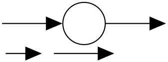
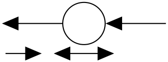


- Game[1]
  - Can you get a ball from node in set X to node in set Y without being blocked by node set Z?
  - Depending on its direction and the previous node, the ball can
    - **Pass through** (from parent to all children, from child to all parents)
    - **Bounce back** (from any parent/child to all parents/children)
    - **Be blocked**

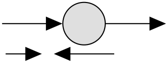
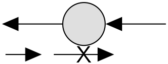
[1] Reference :R.D. Shachter, Bayes-Ball: The Rational Pastime (for Determining Irrelevance and Requisite Information in Belief Networks and Influence Diagrams), UAI'98, 1998

"Bayes Ball" - Recap

- An **unobserved** node $W \notin V$ **passes through** balls from parents and additionally **bounces back** balls from children

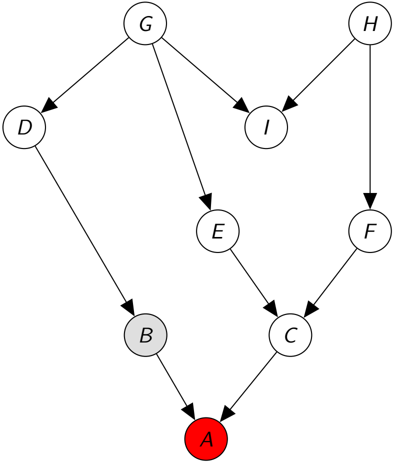
 

- An **observed** node $W \in V$ **bounces back** balls from parents, but **blocks** balls from children

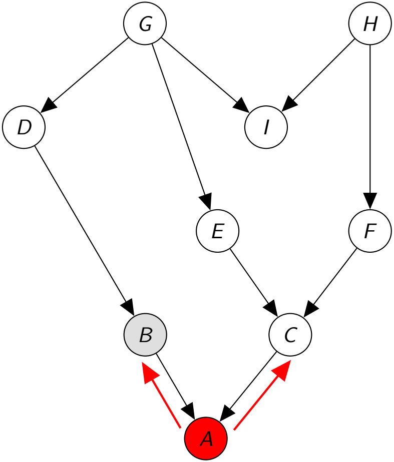
$\Rightarrow$ The Bayes Ball algorithm determines those nodes that are d-separated from the query node.

"Bayes Ball" - Q&A



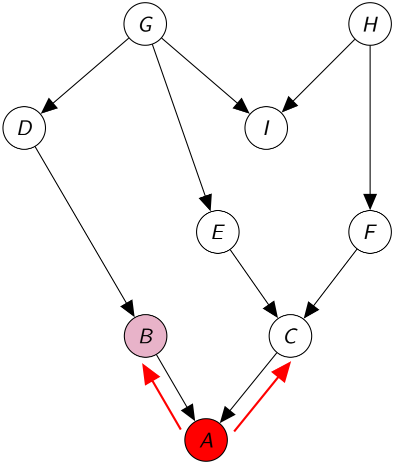
Let's play Bayes Ball...

"Bayes Ball" - Q&A



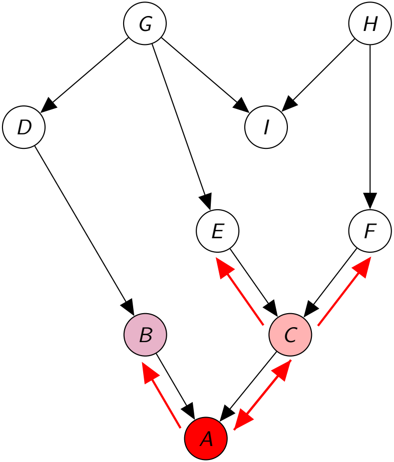
Let's play Bayes Ball...

"Bayes Ball" - Q&A



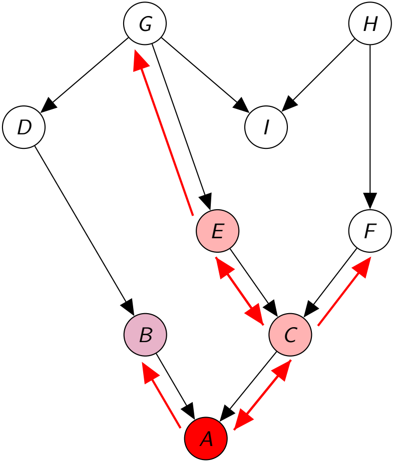
Let's play Bayes Ball...

"Bayes Ball" - Q&A



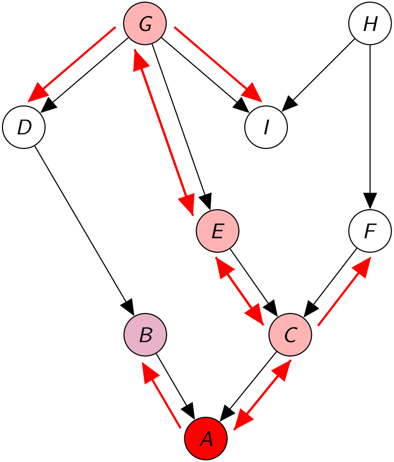
Let's play Bayes Ball...

"Bayes Ball" - Q&A



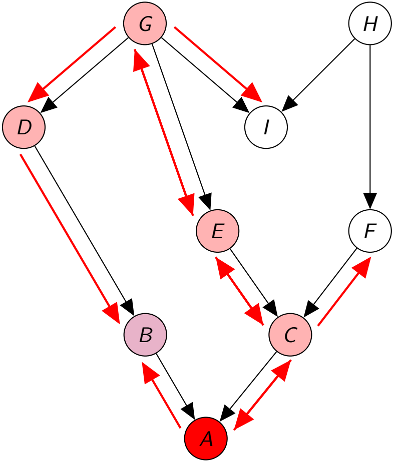
Let's play Bayes Ball...

“Bayes Ball” - Q&A



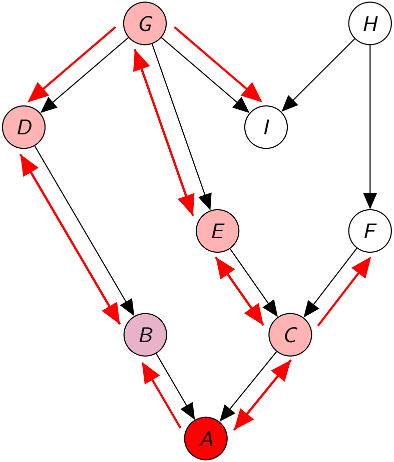
Let’s play Bayes Ball...

"Bayes Ball" - Q&A



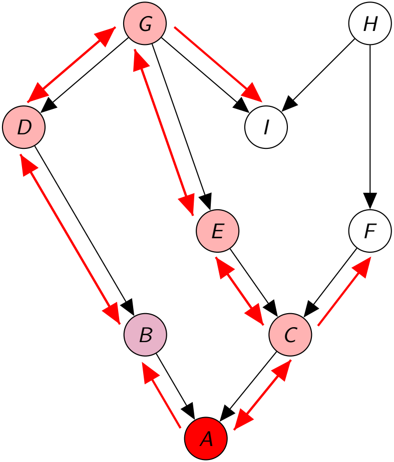
Let's play Bayes Ball...

“Bayes Ball” - Q&A



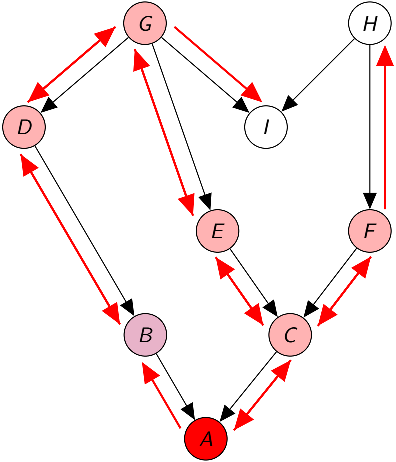
Let’s play Bayes Ball...

"Bayes Ball" - Q&A



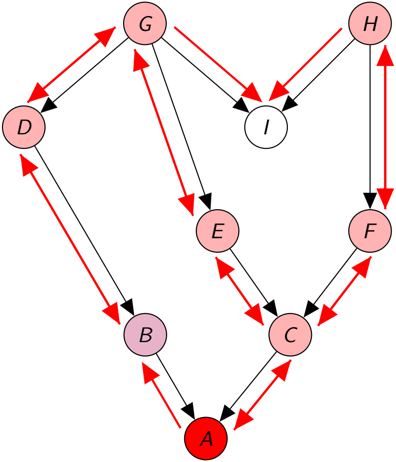
Let's play Bayes Ball...

“Bayes Ball” - Q&A



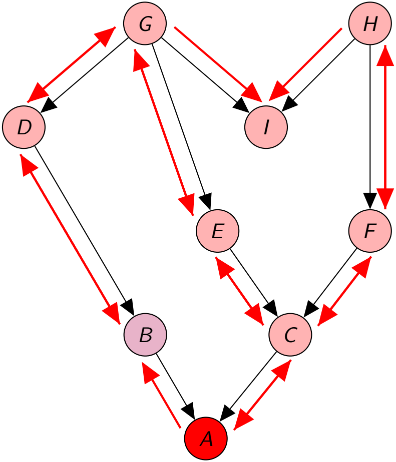
Let’s play Bayes Ball...

"Bayes Ball" - Q&A



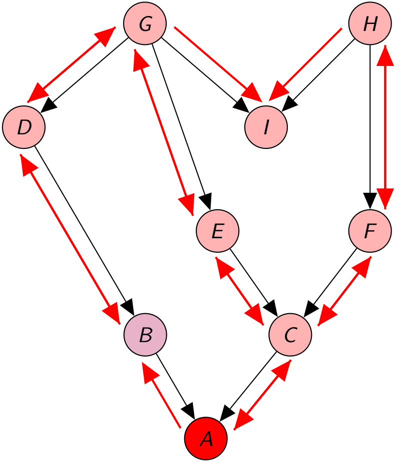
Let's play Bayes Ball...

"Bayes Ball" - Q&A



Let's play Bayes Ball...

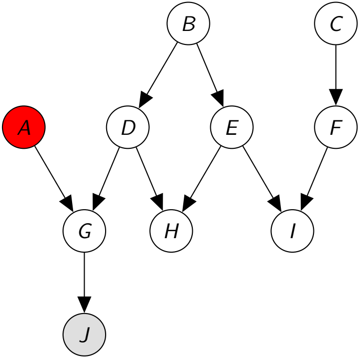
“Bayes Ball” - Q&A



Let’s play Bayes Ball...

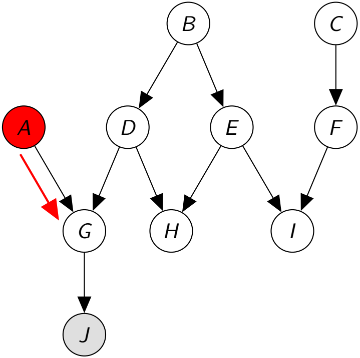
No variable is conditionally independent of A given B

"Bayes Ball" - Q&A



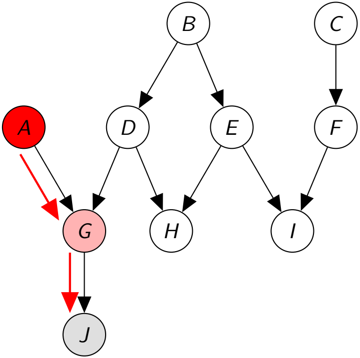
Let's play Bayes Ball...

"Bayes Ball" - Q&A



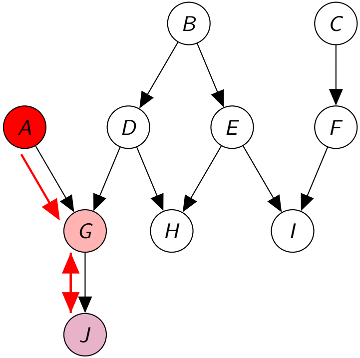
Let's play Bayes Ball...

"Bayes Ball" - Q&A



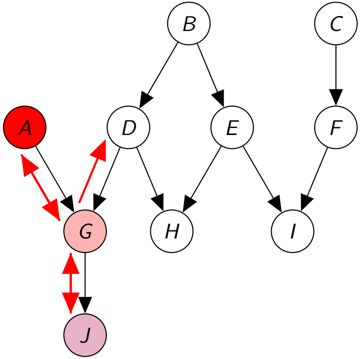
Let's play Bayes Ball...

“Bayes Ball” - Q&A



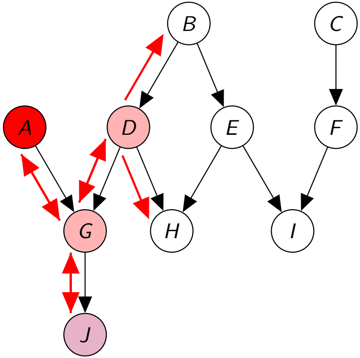
Let’s play Bayes Ball...

"Bayes Ball" - Q&A



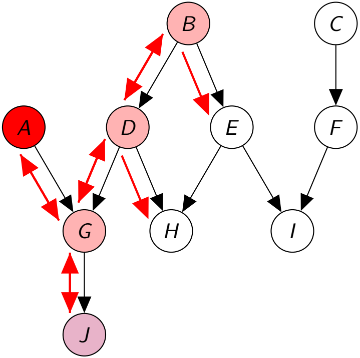
Let's play Bayes Ball...

"Bayes Ball" - Q&A



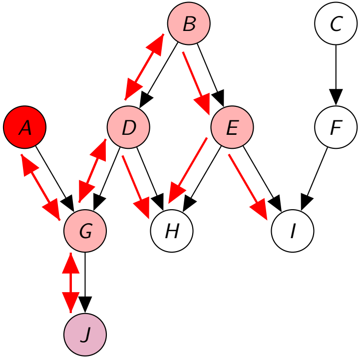
Let's play Bayes Ball...

“Bayes Ball” - Q&A



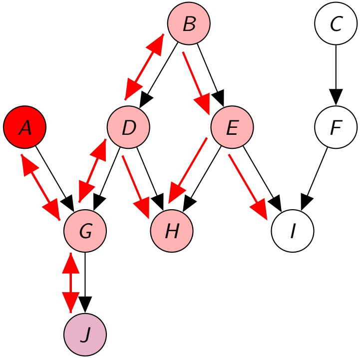
Let’s play Bayes Ball...

"Bayes Ball" - Q&A



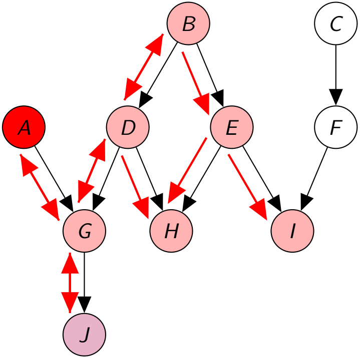
Let's play Bayes Ball...

"Bayes Ball" - Q&A



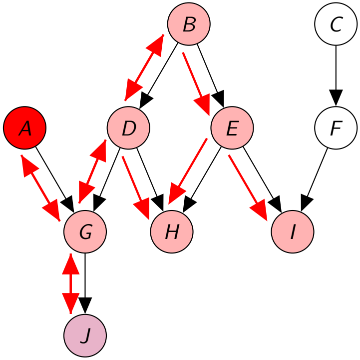
Let's play Bayes Ball...

"Bayes Ball" - Q&A



Let's play Bayes Ball...

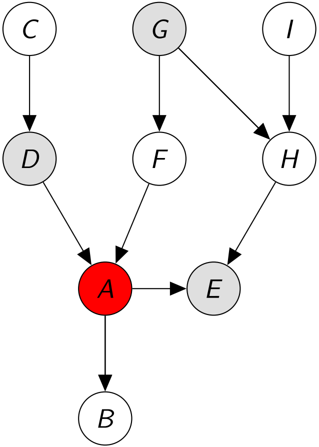
"Bayes Ball" - Q&A



Let's play Bayes Ball...

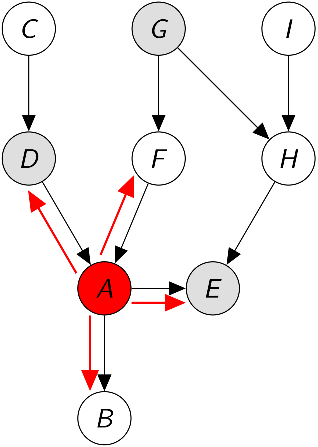
F and C are conditionally independent of A given J

"Bayes Ball" - Q&A



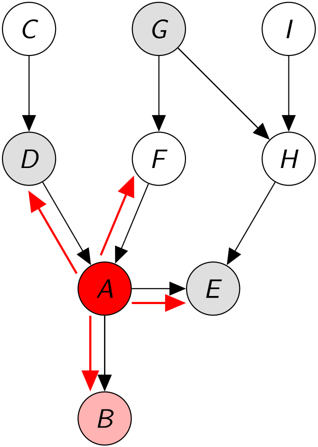
Let's play Bayes Ball...

“Bayes Ball” - Q&A



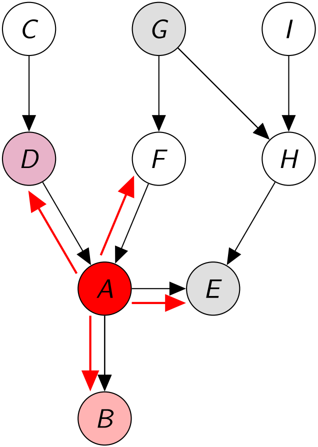
Let’s play Bayes Ball...

"Bayes Ball" - Q&A



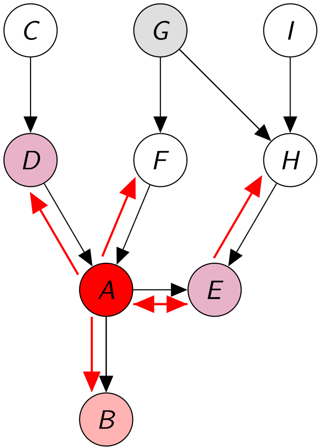
Let's play Bayes Ball...

"Bayes Ball" - Q&A



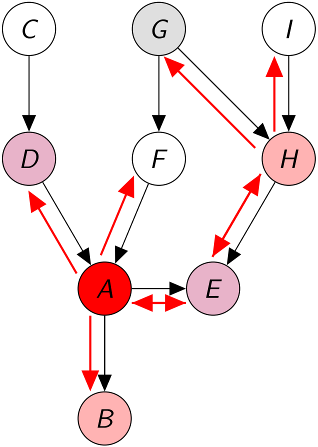
Let's play Bayes Ball...

"Bayes Ball" - Q&A



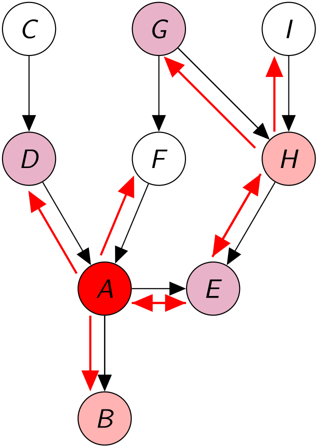
Let's play Bayes Ball...

“Bayes Ball” - Q&A



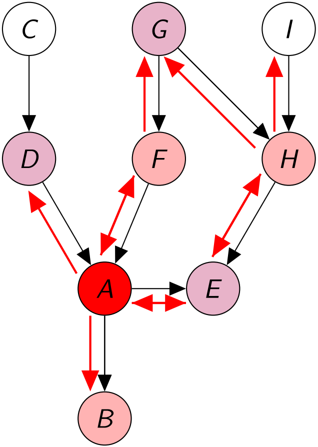
Let’s play Bayes Ball...

"Bayes Ball" - Q&A



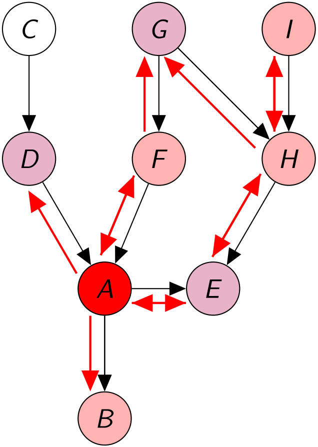
Let's play Bayes Ball...

"Bayes Ball" - Q&A



Let's play Bayes Ball...

“Bayes Ball” - Q&A



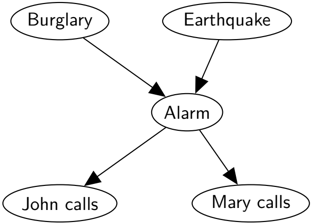
Let’s play Bayes Ball...

C is conditionally independent of A given D, G, E

Content

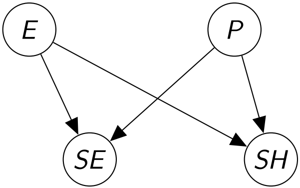
Bayes Net Toolbox - Solution

- If John calls, probability of burglary $= 0.02$
- If there is a burglary, probability that John calls $= 0.85$
- If John & Mary call, and there is no earthquake, probability of burglary $= 0.34$



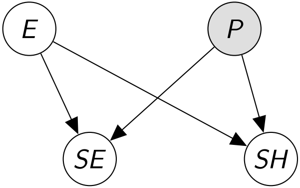
Bayes Net Toolbox - Solution

- If E=True, SE=high and SH=moderate
  - The probability of candidate being in party $A = 0.176471$
  - The probability of candidate being in party $B = 0.823529$
  - The probability of candidate being in party $C = 0.000000$



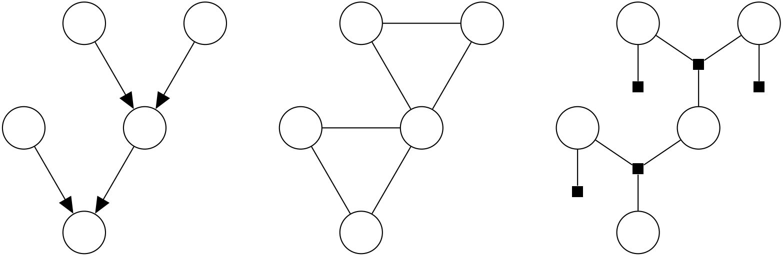
Bayes Net Toolbox - Solution

- If E=True, SE=high and SH=moderate and $p(P = C) = 0$
  - The probability of candidate being in party $A = 0.176471$
  - The probability of candidate being in party $B = 0.823529$
  - The probability of candidate being in party $C = 0.000000$



Content

Factor Graphs - Reminder



- Joint probability
  - Can be expressed as **product of factors**:

$$p(\mathbf{x}) = \prod_{s \in \mathrm{ne}(x)} F_s(x, X_s)$$

  - Factor graphs make this explicit through separate factor nodes
- Converting a directed polytree
  - Conversion to undirected tree creates loops due to moralization!
  - Conversion to a factor graph again results in a tree!

Sum-Product Algorithm - Reminder

- Objectives
  - Efficient, **exact inference** algorithm for finding marginals
- Procedure:
  - **Pick an arbitrary node** as root
  - Compute and propagate messages **from the leaf nodes to the root**, storing received messages at every node
  - Compute and propagate messages **from the root to the leaf nodes**, storing received messages at every node
  - Compute the **product of received messages at each node** for which the marginal is required, and normalize if necessary

$$p(x) \propto \prod_{s \in \mathrm{ne}(x)} \mu_{f_s \to x}(x)$$

- Computational effort
  - Total number of messages $= 2\cdot$ number of graph edges

Sum-Product Algorithm - Reminder

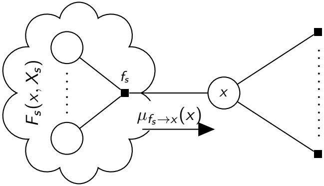
- Two kinds of messages
  - Message from factor node to variable nodes:
    - Sum of factor contributions

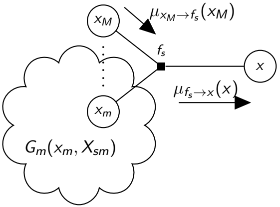
$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$
$$= \sum_{X_s} f_s(x_s) \prod_{m \in \text{ne}(f_s) \setminus x} \mu_{x_m \to f_s}(x_m)$$

  - Message from variable node to factor node:
    - Sum of factor contributions

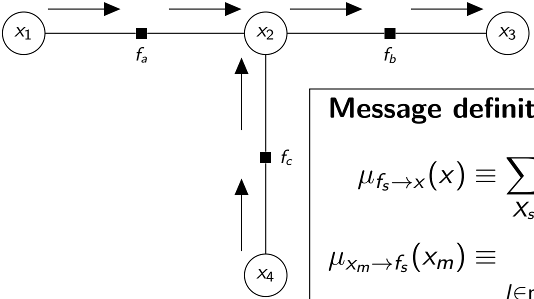
$$\mu_{x_m \to f_s}(x_m) = \prod_{l \in \text{ne}(x_m) \setminus f_s} \mu_{f_l \to x_m}(x_m)$$

$\Rightarrow$ Simple propagation scheme





Sum-Product Algorithm - Reminder



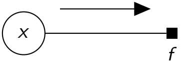
Message definitions:

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} f_s(X_s) \prod_{m \in \text{ne}(f_s) \setminus x} \mu_{x_m \to f_s}(x_m)$$

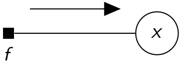
$$\mu_{x_m \to f_s}(x_m) \equiv \prod_{l \in \text{ne}(x_m) \setminus f_s} \mu_{f_l \to x_m}(x_m)$$

Initialization:

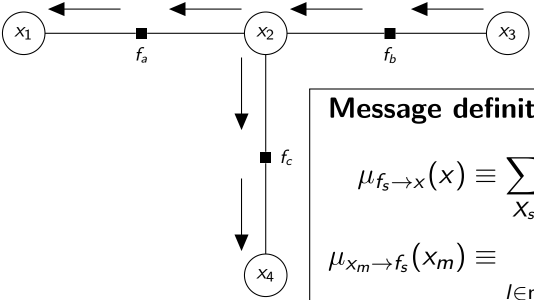
$\mu_{x \to f}(x) = 1$



$\mu_{f \to x}(x) = f(x)$



Sum-Product Algorithm - Reminder



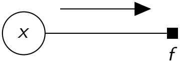
Message definitions:

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} f_s(x_s) \prod_{m \in \mathrm{ne}(f_s) \setminus x} \mu_{x_m \to f_s}(x_m)$$

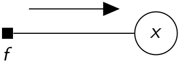
$$\mu_{x_m \to f_s}(x_m) \equiv \prod_{l \in \mathrm{ne}(x_m) \setminus f_s} \mu_{f_l \to x_m}(x_m)$$

Initialization:

$$\mu_{x \to f}(x) = 1$$

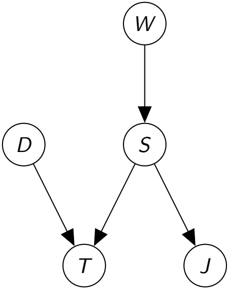


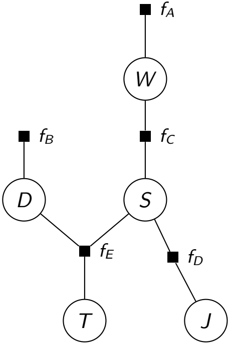
$$\mu_{f \to x}(x) = f(x)$$



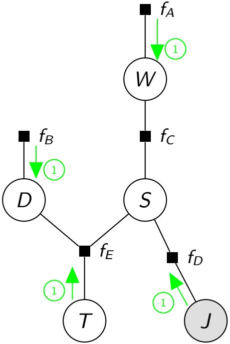
Factor Graph - Solution

Conversion of Bayesian Network to Factor Graph





Sum-Product Algorithm - Solutions



Step 1

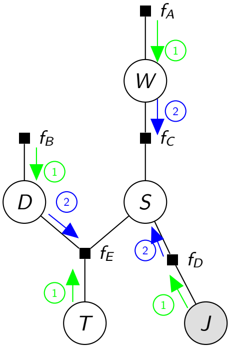
$$\mu_{f_A \to W}(W) = \sum_{\mathrm{ne}(f_A)\setminus\{W\}} f_A(W) = f_A(W)$$

$$\mu_{f_B \to D}(D) = \sum_{\mathrm{ne}(f_B)\setminus\{D\}} f_B(D) = f_B(D)$$

$$\mu_{T \to f_E}(T) = 1$$

$$\mu_{J \to f_D}(J) = 1$$

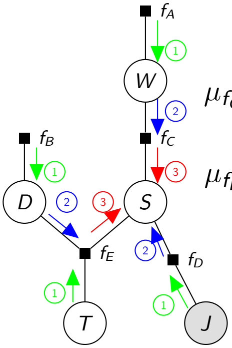
Sum-Product Algorithm - Solutions



Step 2

$$\mu_{W \to f_C}(W) = \mu_{f_A \to W}(W)$$
$$\mu_{D \to f_E}(D) = \mu_{f_B \to D}(D)$$
$$\mu_{f_D \to S}(S) = \mu_{J \to f_D}(J) f_D(S, J = T)$$

Sum-Product Algorithm - Solutions

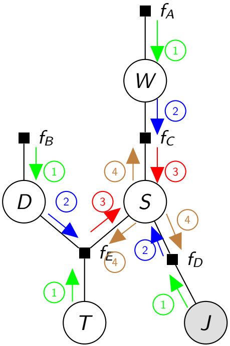


Step 3

$$\mu_{f_C \to S}(S) = \sum_{\mathrm{ne}(f_C)\setminus\{S\}} \mu_{W \to f_C}(W) f_C(S, W)$$

$$\mu_{f_E \to S}(S) = \sum_{\mathrm{ne}(f_E)\setminus\{S\}} \mu_{D \to f_E}(D)\mu_{T \to f_E}(T) f_E(S, D, T)$$

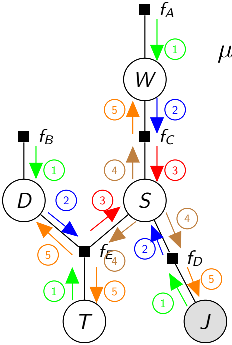
Sum-Product Algorithm - Solutions



Step 4

$$\mu_{S \to f_C}(S) = \mu_{f_D \to S}(S)\mu_{f_E \to S}(S)$$
$$\mu_{S \to f_D}(S) = \mu_{f_C \to S}(S)\mu_{f_E \to S}(S)$$
$$\mu_{S \to f_E}(S) = \mu_{f_C \to S}(S)\mu_{f_D \to S}(S)$$

Sum-Product Algorithm - Solutions



Step 5

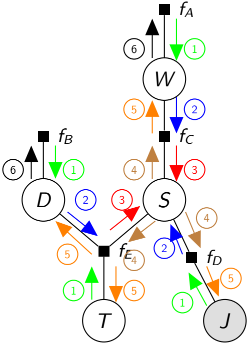
$$\mu_{f_C \to W}(W) = \sum_{\text{ne}(f_C)\setminus\{W\}} \mu_{S \to f_C}(S) f_C(S, W)$$

$$\mu_{f_D \to J}(S) = \sum_{\text{ne}(f_D)\setminus\{J\}} \mu_{S \to f_D}(S) f_D(S, J)$$

$$\mu_{f_E \to T}(T) = \sum_{\text{ne}(f_E)\setminus\{T\}} \mu_{D \to f_E}(S) \mu_{D \to f_E}(D) f_E(D, S, T)$$

$$\mu_{f_E \to D}(D) = \sum_{\text{ne}(f_E)\setminus\{D\}} \mu_{S \to f_E}(S) \mu_{T \to f_E}(T) f_E(D, S, T)$$

Sum-Product Algorithm - Solutions



Step 6

$$\mu_{W\to f_A}(W) = \mu_{f_C\to W}(W)$$
$$\mu_{D\to f_B}(D) = \mu_{f_E\to D}(D)$$

Sum-Product Algorithm - Solutions

- Termination

$$
\begin{aligned}
p(S) &= \mu_{f_D\to S}(S)\mu_{f_C\to S}(S)\mu_{f_E\to S}(S)\\
&= \mu_{f_E\to S}(S) f_D(S, J=T) \sum_{ne(f_E)\setminus\{S\}} \mu_{W\to f_C}(W) f_C(S,W)\cdot\\
&\quad\cdot \sum_{ne(f_C)\setminus\{S\}} \mu_{D\to f_E}(D)\mu_{T\to f_E}(T) f_E(S,D,T)\\
&= f_D(S,J=T) \sum_{ne(f_C)\setminus S} f_A(W) f_C(S,W) \sum_{ne(f_E)\setminus S} f_B(D) f_E(S,D,T)
\end{aligned}
$$

- Results

$$p(S=T) = 0.7692 \qquad p(S=F) = 0.2308$$

- If we have no observation we have

$$p(S=T) = 0.4 \qquad p(S=F) = 0.6$$