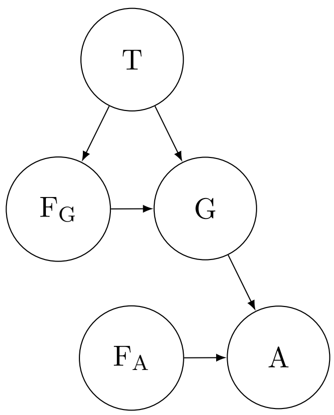
# COMP 424 Assignment 3 Solutions

Ariella Smofsky

April 4, 2017

**Q1** (a)




(b) The above network *is not* a polytree because if we ignore the directions on edges the network has an undirected cycle between nodes T, $F_G$, and G.

(c)

| | $F_G$ = True<br>T = High | $F_G$ = True<br>T = Normal | $F_G$ = False<br>T = High | $F_G$ = False<br>T = Normal |
|---|---|---|---|---|
| G = High | y | 1 - y | x | 1 - x |
| G = Normal | 1 - y | y | 1 - x | x |

(d)

| | $F_A$ = True<br>G = High | $F_A$ = True<br>G = Normal | $F_A$ = False<br>G = High | $F_A$ = False<br>G = Normal |
|---|---|---|---|---|
| A = True | 0 | 0 | 1 | 0 |
| A = False | 1 | 1 | 0 | 1 |

(e) We are interested in calculating

$$P(T|A, \neg F_A, \neg F_G).$$

Looking at the table from part (d), we can use the information A = True and $F_A$ = False to conclude that G = High. Thus, we are interested in calculating

$$P(T|A, \neg F_A, \neg F_G, G).$$

Use the definition of conditional probability to obtain

$$P(T|A, \neg F_A, \neg F_G, G) = \frac{P(T, A, \neg F_A, \neg F_G, G)}{P(A, \neg F_A, \neg F_G, G)}$$

Use chain rule

$$= \frac{P(A|T, \neg F_A, \neg F_G, G) \times P(\neg F_A|T, \neg F_G, G) \times P(G|\neg F_G, T) \times P(\neg F_G|T) \times P(T)}{P(A|\neg F_A, G, \neg F_G) \times P(\neg F_A|G, \neg F_G) \times P(\neg F_G|G) \times P(G)}$$

By conditional independence

$$= \frac{P(A|\neg F_A, G) \times P(\neg F_A) \times P(G|\neg F_G, T) \times P(\neg F_G|T) \times P(T)}{P(A|\neg F_A, G) \times P(\neg F_A) \times P(\neg F_G) \times P(G)}$$

Cancel common terms

$$= \frac{P(G|\neg F_G, T) \times P(\neg F_G|T) \times P(T)}{P(\neg F_G, G)}$$

Marginalize the denominator over hidden variable T

$$= \frac{P(G|\neg F_G, T) \times P(\neg F_G|T) \times P(T)}{P(\neg F_G, G, T) + P(\neg F_G, G, \neg T)}$$

Use chain rule in the denominator

$$= \frac{P(G|\neg F_G, T) \times P(\neg F_G|T) \times P(T)}{P(G|\neg F_G, T)P(\neg F_G|T)P(T) + P(G|\neg F_G, \neg T)P(\neg F_G|\neg T)P(\neg T)}$$

To simplify notation let $P(T) = m$, $P(\neg T) = 1 - m$, $P(\neg F_G|T) = n$, $P(\neg F_G|\neg T) = q$. Use the table from part (c) to get the final desired expression

$$= \frac{xnm}{xnm + (1-x)q(1-m)}$$

**Q2** (a) We want to calculate

$$P(g, s).$$

Marginalize over hidden variables

$$P(g, s) = \sum_{Q,H,E} P(Q, H, E, g, s)$$

Determine the joint probability using the structure of the Bayesian network

$$= \sum_{Q,H,E} P(g|Q,H)P(Q)P(H|Q,s)P(E|s)P(s)$$

Factor out the term $P(s)$ and expand the sum for hidden variable $E$

$$= P(s)[(P(e|s) + P(\neg e|s))(\sum_{Q,H} P(g|Q,H)P(Q)P(H|Q,s))]$$

Note that $P(e|s) + P(\neg(e)|s) = 1$. Thus we have

$$= 0.7 \sum_{Q,H} P(g|Q,H)P(Q)P(H|Q,s)$$

Expand the sum for hidden variables $Q$ and $H$

$$\begin{aligned} = 0.7[&P(g|Q,H)P(Q)P(H|Q,s) \\ &+P(g|Q,\neg H)P(Q)P(\neg H|Q,s) \\ &+P(g|\neg Q,H)P(\neg Q)P(H|\neg Q,s) \\ &+P(g|\neg Q,\neg H)P(\neg Q)P(\neg H|\neg Q,s)] \end{aligned}$$

Finally, plug in given values

$$\begin{aligned} = 0.7[&0.75 \times 0.8 \times 0.9 \\ &+0.4 \times 0.8 \times 0.1 \\ &+0.6 \times 0.2 \times 0.15 \\ &+0.3 \times 0.2 \times 0.85] \end{aligned}$$

and we get the final answer

$$= 0.4487$$

(b) We want to calculate

$$P(q, \neg e).$$

Marginalize over hidden variables

$$P(q, \neg e) = \sum_{H,G,S} P(q, H, \neg e, G, S)$$

Determine the joint probability using the structure of the Bayesian network

$$= \sum_{H,G,S} P(G|q,H)P(q)P(H|q,S)P(\neg e|S)P(S)$$

Factor out the term $P(q)$ and expand the sum for hidden variable $G$

$$= P(q)[(P(G|q,H) + P(\neg G|q,H)) \sum_{H,S} P(H|q,S)P(\neg e|S)P(S)]$$

Note that $P(G|q,H) + P(\neg G|q,H) = 1$. Thus we have

$$= 0.8 \sum_{H,S} P(H|q,S)P(\neg e|S)P(S)$$

Expand the sum for hidden variables $H$ and $S$

$$\begin{aligned} = 0.8[&P(h|q,s)P(\neg e|s)P(s) \\ &+P(\neg h|q,s)P(\neg e|s)P(s) \\ &+P(h|q,\neg s)P(\neg e|\neg s)P(\neg s) \\ &+P(\neg h|q,\neg s)P(\neg e|\neg s)P(\neg s)] \end{aligned}$$

Finally, plug in given values

$$\begin{aligned} = 0.8[&0.9 \times 0.3 \times 0.7 \\ &+0.1 \times 0.3 \times 0.7 \\ &+0.85 \times 0.5 \times 0.3 \\ &+0.15 \times 0.5 \times 0.3] \end{aligned}$$

and we get the final answer

$$= 0.288$$

(c) We want to calculate

$$P(g|e).$$

Use the definition of conditional probability

$$P(g|e) = \frac{P(g,e)}{P(e)}$$

Marginalize over hidden variables

$$= \frac{\sum_{Q,H,S} P(Q,H,S,g,e)}{\sum_{Q,H,S,G} P(Q,H,S,G,e)}$$

Determine the joint probability using the structure of the Bayesian network

$$= \frac{\sum_{Q,H,S} P(g|Q,H)P(Q)P(H|Q,S)P(S)P(e|S)}{\sum_{Q,H,S,G} P(G|Q,H)P(Q)P(H|Q,S)P(S)P(e|S)}$$

Simplify the denominator by using Bayes Ball

$$P(e) = \sum_{S} P(S)P(e|S) = 0.7 \times 0.7 + 0.3 \times 0.5 = 0.64$$

Expand the sum for hidden variables $Q$, $H$, and $S$

$$\begin{aligned}
&= \frac{1}{0.64}[P(g|q,h)P(q)P(h|q,s)P(s)P(e|s) \\
&\quad +P(g|q,h)P(q)P(h|q,\neg s)P(\neg s)P(e|\neg s) \\
&\quad +P(g|q,\neg h)P(q)P(\neg h|q,s)P(s)P(e|s) \\
&\quad +P(g|q,\neg h)P(q)P(\neg h|q,\neg s)P(\neg s)P(e|\neg s) \\
&\quad +P(g|\neg q,h)P(\neg q)P(h|\neg q,s)P(s)P(e|s) \\
&\quad +P(g|\neg q,h)P(\neg q)P(h|\neg q,\neg s)P(\neg s)P(e|\neg s) \\
&\quad +P(g|\neg q,\neg h)P(\neg q)P(\neg h|\neg q,s)P(s)P(e|s) \\
&\quad +P(g|\neg q,\neg h)P(\neg q)P(\neg h|\neg q,\neg s)P(\neg s)P(e|\neg s)]
\end{aligned}$$

Finally, plug in given values

$$\begin{aligned}
&= \frac{1}{0.64}[0.75 \times 0.8 \times 0.9 \times 0.7 \times 0.7 \\
&\quad +0.75 \times 0.8 \times 0.85 \times 0.3 \times 0.5 \\
&\quad +0.4 \times 0.8 \times 0.1 \times 0.7 \times 0.7 \\
&\quad +0.4 \times 0.8 \times 0.15 \times 0.3 \times 0.5 \\
&\quad +0.6 \times 0.2 \times 0.15 \times 0.7 \times 0.7 \\
&\quad +0.6 \times 0.2 \times 0.3 \times 0.3 \times 0.5 \\
&\quad +0.3 \times 0.2 \times 0.85 \times 0.7 \times 0.7 \\
&\quad +0.3 \times 0.2 \times 0.7 \times 0.3 \times 0.5]
\end{aligned}$$

and we get the final answer

$$= 0.6398$$

**Q3** We want to compute the MAP of

$$P(E|h).$$

Use the definition of conditional probability. Note that $\alpha$ is the normalizing constant

$$P(E|h) = \frac{P(E,h)}{P(h)} = \alpha P(E,h)$$

Marginalize over hidden variables $Q$, $S$, and $G$

$$= \alpha \sum_{Q,S,G} P(Q,h,S,G,E)$$

Determine the joint probability using the structure of the Bayesian network

$$= \alpha \sum_{Q,S,G} P(G|Q,h)P(Q)P(h|Q,S)P(S)P(E|S)$$

Apply variable elimination in the order $G, S, Q, E$

$$= \alpha \underbrace{\sum_q P(q)}_{f_1(Q)} \sum_s \underbrace{P(s)}_{f_2(S)} \underbrace{P(E|s)}_{f_3(E,S)} \underbrace{P(h|q,s)}_{f_4(Q,S)} \underbrace{\sum_g P(g|q,h)}_{\text{eliminate } G \text{ first}}$$

Apply variable elimination to $G$

$$= \alpha \sum_q f_1(Q) \sum_s f_2(S) f_3(E,S) f_4(Q,S)$$

Apply variable elimination to $S$

$$= \alpha \sum_q f_1(Q) f_5(E,Q)$$

Apply variable elimination to $Q$

$$= \alpha f_6(E)$$

We are now done elimination. The intermediate factors created are

$$\begin{aligned} f_5(E,Q) &= f_2(s)f_3(E,s)f_4(Q,s) + f_2(\neg s)f_3(E,\neg s)f_4(Q,\neg s) \\ f_6(E) &= f_1(q)f_5(E,q) + f_1(\neg q)f_5(E,\neg q) \end{aligned}$$

Now to calculate the MAP of $P(E|h)$

$$\begin{aligned} f_2(S) &= \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \\ f_3(E,s) &= \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \\ f_3(E,\neg s) &= \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \\ f_4(Q,s) &= \begin{bmatrix} 0.9 \\ 0.15 \end{bmatrix} \\ f_4(Q,\neg s) &= \begin{bmatrix} 0.85 \\ 0.3 \end{bmatrix} \end{aligned}$$

$$
\begin{aligned}
f_5(E,Q) &= 0.7 \begin{bmatrix}0.7\\0.3\end{bmatrix} \begin{bmatrix}0.9\\0.15\end{bmatrix} + 0.3 \begin{bmatrix}0.5\\0.5\end{bmatrix} \begin{bmatrix}0.85\\0.3\end{bmatrix} \\
&= 0.7 \begin{bmatrix}0.63 & 0.105\\0.27 & 0.045\end{bmatrix} + 0.3 \begin{bmatrix}0.425 & 0.15\\0.425 & 0.15\end{bmatrix} \\
&= \begin{bmatrix}0.5685 & 0.1185\\0.3165 & 0.0765\end{bmatrix} \\
f_6(E) &= f_1(q)f_5(E,q) + f_1(\neg q)f_5(E,\neg q) \\
&= 0.8 \begin{bmatrix}0.5685\\0.3165\end{bmatrix} + 0.2 \begin{bmatrix}0.1185\\0.0765\end{bmatrix} \\
&= \begin{bmatrix}0.4785\\0.2685\end{bmatrix}
\end{aligned}
$$

Calculate the value of normalizing constant $\alpha$

$$
\begin{aligned}
P(e|h) + P(\neg e|h) = \alpha f_6(e) + \alpha f_6(\neg e) = 1 \\
\Rightarrow \alpha = \frac{1}{f_6(e) + f_6(\neg e)} = \frac{1}{0.4785 + 0.2685} = 1.3387
\end{aligned}
$$

Thus

$$
P(E|h) = \alpha f_6(E) = 1.3387 \begin{bmatrix}0.4785\\0.2685\end{bmatrix} = \begin{bmatrix}0.6406\\0.3594\end{bmatrix}
$$

Clearly, $0.6406 > 0.3594$. So we can finally conclude the MAP of $P(E|h)$ is

$$
E = e
$$

**Q4** (a) i. The parameters that must be learned are

$$
\begin{aligned}
\theta_a &= P(a) \\
\theta_{b|a} &= P(b|a) \\
\theta_{b|\neg a} &= P(b|\neg a) \\
\theta_{c|a} &= P(c|a) \\
\theta_{c|\neg a} &= P(c|\neg a) \\
\theta_{d|b,c} &= P(d|b,c) \\
\theta_{d|b,\neg c} &= P(d|b,\neg c) \\
\theta_{d|\neg b,c} &= P(d|\neg b,c) \\
\theta_{d|\neg b,\neg c} &= P(d|\neg b,\neg c)
\end{aligned}
$$

ii. Calculate the MLE of each parameter

$$
\begin{aligned}
\theta_a &= \frac{35}{135} \\
\theta_{b|a} &= \frac{20}{35} \\
\theta_{b|\neg a} &= \frac{66}{100} \\
\theta_{c|a} &= \frac{33}{35} \\
\theta_{c|\neg a} &= \frac{90}{100} \\
\theta_{d|b,c} &= \frac{7}{80} \\
\theta_{d|b,\neg c} &= \frac{3}{6} \\
\theta_{d|\neg b,c} &= \frac{22}{43} \\
\theta_{d|\neg b,\neg c} &= \frac{4}{6}
\end{aligned}
$$

iii. Calculate the MAP of each parameter by adding 1 to each numerator and 2 to each denominator because all variables are Bernoulli

$$
\begin{aligned}
\theta_a &= \frac{36}{137} \\
\theta_{b|a} &= \frac{21}{37} \\
\theta_{b|\neg a} &= \frac{67}{102} \\
\theta_{c|a} &= \frac{34}{37} \\
\theta_{c|\neg a} &= \frac{91}{102} \\
\theta_{d|b,c} &= \frac{8}{82} \\
\theta_{d|b,\neg c} &= \frac{4}{8} \\
\theta_{d|\neg b,c} &= \frac{23}{45} \\
\theta_{d|\neg b,\neg c} &= \frac{5}{8}
\end{aligned}
$$

(b) i. Compute the first E-step by calculating

$$P(B|a, \neg c, \neg d),\ P(D|a, b, c)$$

Calculate the first probability

$$P(B|a, \neg c, \neg d) = \frac{P(a, B, \neg c, \neg d)}{P(a, \neg c, \neg d)} = \alpha P(a, B, \neg c, \neg d) = \alpha P(a)P(B|a)P(\neg c|a)P(\neg d|B, \neg c)$$

Incorporate constants $P(a)$ and $P(\neg c|a)$ into new factorization constant $\beta$

$$= \beta P(B|a)P(\neg d|B, \neg c)$$

Let $B = b$

$$\beta P(b|a)P(\neg d|b, \neg c) = \beta(\frac{20}{35})(\frac{3}{6})$$

Let $B = \neg b$

$$\beta P(\neg b|a)P(\neg d|\neg b, \neg c) = \beta(\frac{15}{35})(\frac{2}{6})$$

Normalize to get $\beta = 2.33$. Plug in this value to get desired probabilities

$$P(b|a, \neg c, \neg d) = \frac{2}{3} = 0.666$$
$$P(\neg b|a, \neg c, \neg d) = \frac{1}{3} = 0.333$$

Calculate the second probability

$$P(D|a, b, c) = \alpha P(a, b, c, D) = \alpha P(a)P(b|a)P(c|a)P(D|b, c)$$

Incoporate constants $P(a), P(b|a), P(c|a)$ into new factorization constant $\beta$

$$= \beta P(D|b, c)$$

Let $D = d$

$$\beta P(d|b, c) = \beta\frac{7}{80}$$

Let $D = \neg d$

$$\beta P(d|b, c) = \beta\frac{73}{80}$$

Clearly the normalizing constant $\beta = 1$ and we have

$$P(d|a, b, c) = \frac{7}{80}$$
$$P(\neg d|a, b, c) = \frac{73}{80}$$

ii. Compute the first M-step by calculating

$$P(b|a) = \frac{20 + 1 + \frac{2}{3}}{35 + 2} = 0.5856$$
$$P(D|b,c) = \frac{(7 + 0.875)}{80 + 1} = 0.0875$$

Note the value of $P(D|b,c)$ did not change during the first M-Step. Now update all parameters

$$\begin{aligned}
\theta_a &= \frac{35 + 2}{135 + 2} = \frac{37}{137} \\
\theta_{b|a} &= 0.558559 \\
\theta_{b|\neg a} &= \text{unchanged} \\
\theta_{c|a} &= \frac{33 + 1}{35 + 2} = \frac{34}{37} \\
\theta_{c|\neg a} &= \text{unchanged} \\
\theta_{d|b,c} &= 0.0875 \\
\theta_{d|b,\neg c} &= \frac{3}{6 + \frac{2}{3}} = 0.45 \\
\theta_{d|\neg b,c} &= \text{unchanged} \\
\theta_{d|\neg b,\neg c} &= \frac{4}{6 + \frac{1}{3}} = 0.6316
\end{aligned}$$

iii. Compute the second E-step by updating

$$P(B|a, \neg c, \neg d)$$

Let $B = b$

$$\begin{aligned}
P(b|a, \neg c, \neg d) &= \alpha P(b|a) P(\neg d|b, \neg c) \\
&= \alpha(0.5856)(0.55) \\
&= \alpha(0.32208)
\end{aligned}$$

Let $B = \neg b$

$$\begin{aligned}
P(\neg b|a, \neg c, \neg d) &= \alpha P(\neg b|a) P(\neg d|\neg b, \neg c) \\
&= \alpha(0.4144)(0.3683) \\
&= \alpha(0.1526)
\end{aligned}$$

Finally we have

$$P(b|a, \neg c, \neg d) = 0.6785$$
$$P(\neg b|a, \neg c, \neg d) = 0.3215$$

The second probability $P(D|a, b, c)$ remains unchanged because the first M-step did not alter the value of $P(D|b, c)$.