TU MÜNCHEN
GRUNDLAGEN DER KÜNSTLICHEN INTELLIGENZ
WS 2019/2020

Solutions for exercise 8 11 January 2020

BAYESIAN NETWORKS

Problem 8.1:

a. ii, iii, iv and v.
For v., consider the Markov blanket of D.
Remember that a node in the Bayesian network is conditionally independent from all other nodes given its parents, children and children's parents (this evidence is the **Markov blanket** of the node).

b. $\mathbf{P}(D,TP,CC,UP,PP,W) = \mathbf{P}(D)\mathbf{P}(TP)\mathbf{P}(CC)\mathbf{P}(UP|D,TP)\mathbf{P}(PP|CC,UP)\mathbf{P}(W|CC,UP,PP)$.

c. To calculate these values we can directly use the equation derived in **b.**. In fact, we have to evaluate it using the requested values of the random variables:
$P(\neg d,tp,cc,\neg up,pp,w) = P(\neg d)P(tp)P(cc)P(\neg up|\neg d,tp)P(pp|cc,\neg up)P(w|cc,\neg up,pp) =$
$= 0.75 \cdot 0.35 \cdot 0.3 \cdot 0.9 \cdot 0.9 \cdot 0.9 = 0.05740875$.
$P(\neg d,tp,cc,\neg up,\neg pp,w) = P(\neg d)P(tp)P(cc)P(\neg up|\neg d,tp)P(\neg pp|cc,\neg up)P(w|cc,\neg up,\neg pp) =$
$= 0.75 \cdot 0.35 \cdot 0.3 \cdot 0.9 \cdot 0.1 \cdot 0.5 = 0.00354375$.

d. The required probability is $P(w|tp,\neg d,pp)$.
To calculate this we can use enumeration:
$\mathbf{P}(W|tp,\neg d,pp) = \alpha P(\neg d)P(tp)\sum_{UP}\mathbf{P}(UP|\neg d,tp)\sum_{CC}\mathbf{P}(CC)\mathbf{P}(pp|CC,UP)\mathbf{P}(W|CC,UP,pp)$.

Summing over non-given variables

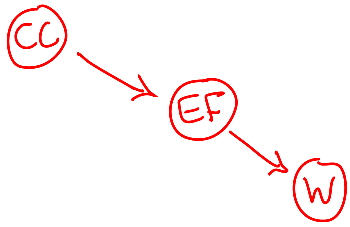
$$\begin{aligned}
P(w|tp,\neg d,pp) = \alpha P(\neg d)P(tp)\Big\{&P(up|\neg d,tp)\big[P(cc)P(pp|cc,up)P(w|cc,up,pp)\\
&+P(\neg cc)P(pp|\neg cc,up)P(w|\neg cc,up,pp)\big]\\
&+P(\neg up|\neg d,tp)\big[P(cc)P(pp|cc,\neg up)P(w|cc,\neg up,pp)\\
&+P(\neg cc)P(pp|\neg cc,\neg up)P(w|\neg cc,\neg up,pp)\big]\Big\}\\
= \alpha \cdot 0.75 \cdot 0.35\Big\{&0.1\big[0.3\cdot 0.5\cdot 0.4+0.7\cdot 0.05\cdot 0.05\big]+0.9\big[0.3\cdot 0.9\cdot 0.9+0.7\cdot 0.1\cdot 0.15\big]\Big\}\\
\approx \alpha \cdot 0.0615,&
\end{aligned}$$

NB In enumeration, normalization 'α' is calculated 2 times $\langle w, \neg w$ since we are summing over other variables

$$\begin{aligned}
P(\neg w|tp,\neg d,pp) = \alpha P(\neg d)P(tp)\Big\{&P(up|\neg d,tp)\big[P(cc)P(pp|cc,up)P(\neg w|cc,up,pp)\\
&+P(\neg cc)P(pp|\neg cc,up)P(\neg w|\neg cc,up,pp)\big]\\
&+P(\neg up|\neg d,tp)\big[P(cc)P(pp|cc,\neg up)P(\neg w|cc,\neg up,pp)\\
&+P(\neg cc)P(pp|\neg cc,\neg up)P(\neg w|\neg cc,\neg up,pp)\big]\Big\}\\
= \alpha \cdot 0.75 \cdot 0.35\Big\{&0.1\big[0.3\cdot 0.5\cdot 0.6+0.7\cdot 0.05\cdot 0.95\big]+0.9\big[0.3\cdot 0.9\cdot 0.1+0.7\cdot 0.1\cdot 0.85\big]\Big\}\\
\approx \alpha \cdot 0.0237,&
\end{aligned}$$

$\alpha = \frac{1}{0.0615+0.0237} \approx 11.7371$,
$P(w|tp,\neg d,pp) \approx 11.7371 \cdot 0.0615 \approx 0.7218$.

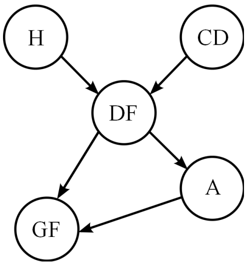
e. The random variable EF has a direct influence on the race. We can reasonably assume that when there is an engine failure during the race, the pilot does not win. Another reasonable assumption is that a competitive car is also more reliable with a direct influence on the random variable EF. In conclusion, we introduce the node EF as a parent of W and child of CC.



Important

Problem 8.2:

a. The corresponding network is the following:



b. Using enumeration:

$$\mathbf{P}(CD|\neg a, gf) = \alpha \mathbf{P}(CD) \sum_{DF} \mathbf{P}(gf|\neg a, DF)\mathbf{P}(\neg a|DF) \sum_{H} \mathbf{P}(DF|CD,H)\mathbf{P}(H).$$

$$\begin{aligned}
P(cd|\neg a, gf) &= \alpha P(cd) \Big\{ P(gf|df,\neg a) P(\neg a|df) \big[P(df|cd,h)P(h) + P(df|cd,\neg h)P(\neg h) \big] \\
&\quad + P(gf|\neg df,\neg a) P(\neg a|\neg df) \big[P(\neg df|cd,h)P(h) + P(\neg df|cd,\neg h)P(\neg h) \big] \Big\} \\
&= \alpha \cdot 0.6 \Big\{ 0.4 \cdot 0.3 \big[0.15 \cdot 0.5 + 0.01 \cdot 0.5 \big] + 0.05 \cdot 0.75 \big[0.85 \cdot 0.5 + 0.99 \cdot 0.5 \big] \Big\} \\
&\approx \alpha \cdot 0.0265,
\end{aligned}$$

$$\begin{aligned}
P(\neg cd|\neg a, gf) &= \alpha P(\neg cd) \Big\{ P(gf|df,\neg a) P(\neg a|df) \big[P(df|\neg cd,h)P(h) + P(df|\neg cd,\neg h)P(\neg h) \big] \\
&\quad + P(gf|\neg df,\neg a) P(\neg a|\neg df) \big[P(\neg df|\neg cd,h)P(h) + P(\neg df|\neg cd,\neg h)P(\neg h) \big] \Big\} \\
&= \alpha \cdot 0.4 \Big\{ 0.4 \cdot 0.3 \big[0.99 \cdot 0.5 + 0.1 \cdot 0.5 \big] + 0.05 \cdot 0.75 \big[0.01 \cdot 0.5 + 0.9 \cdot 0.5 \big] \Big\} \\
&\approx \alpha \cdot 0.0330,
\end{aligned}$$

$$\alpha = \frac{1}{0.0265 + 0.0330} \approx 16.8067,$$

$$\mathbf{P}(CD|\neg a, gf) \approx 16.8067 \cdot \begin{bmatrix} 0.0265 \\ 0.0330 \end{bmatrix} \approx \begin{bmatrix} 0.4454 \\ 0.5546 \end{bmatrix}.$$

c. Using variable elimination:

$$\mathbf{P}(CD|\neg a, gf) = \alpha \underbrace{\mathbf{P}(CD)}_{\mathbf{f}_1(CD)} \sum_{DF} \underbrace{\mathbf{P}(gf|\neg a, DF)}_{\mathbf{f}_2(DF)} \underbrace{\mathbf{P}(\neg a|DF)}_{\mathbf{f}_3(DF)} \sum_{H} \underbrace{\mathbf{P}(DF|CD,H)}_{\mathbf{f}_4(DF,CD,H)} \underbrace{\mathbf{P}(H)}_{\mathbf{f}_5(H)}.$$

In the following notation we represent a $2 \times 2 \times 2$ matrix as $\{ [2 \times 2]\ [2 \times 2] \}$.

$$\begin{aligned}
\mathbf{f}_4(DF, CD, H) &= \Big\{ \ \mathbf{P}(DF|CD,h) \quad \mathbf{P}(DF|CD,\neg h) \ \Big\} \\
&= \left\{ \begin{bmatrix} P(df|cd,h) & P(df|\neg cd,h) \\ P(\neg df|cd,h) & P(\neg df|\neg cd,h) \end{bmatrix} \begin{bmatrix} P(df|cd,\neg h) & P(df|\neg cd,\neg h) \\ P(\neg df|cd,\neg h) & P(\neg df|\neg cd,\neg h) \end{bmatrix} \right\} \\
&= \left\{ \begin{bmatrix} 0.15 & 0.99 \\ 0.85 & 0.01 \end{bmatrix} \begin{bmatrix} 0.01 & 0.1 \\ 0.99 & 0.9 \end{bmatrix} \right\},
\end{aligned}$$

$$\mathbf{f}_1(CD) = \begin{bmatrix} P(cd) & P(\neg cd) \end{bmatrix} = \begin{bmatrix} 0.6 & 0.4 \end{bmatrix}, \quad \mathbf{f}_2(DF) = \begin{bmatrix} P(gf|\neg a, df) \\ P(gf|\neg a, \neg df) \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.05 \end{bmatrix},$$

$$\mathbf{f_3}(DF) = \begin{bmatrix} P(\neg a|df) \\ P(\neg a|\neg df) \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.75 \end{bmatrix}, \ \mathbf{f_5}(H) = \begin{bmatrix} P(h) & P(\neg h) \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}.$$

After defining the factors, we can write:
$\mathbf{P}(CD|\neg a, gf) = \alpha \mathbf{f_1}(CD) \times \sum_{DF} \mathbf{f_2}(DF) \times \mathbf{f_3}(DF) \times \sum_{H} \mathbf{f_4}(DF,CD,H) \times \mathbf{f_5}(H)$.
In the following we use the pointwise product and we sum out variables.

$$\mathbf{f_4}(DF,CD,H) \times \mathbf{f_5}(H) = \left\{ \begin{bmatrix} 0.075 & 0.495 \\ 0.425 & 0.005 \end{bmatrix} \begin{bmatrix} 0.005 & 0.050 \\ 0.495 & 0.450 \end{bmatrix} \right\},$$

$$\mathbf{f_6}(DF,CD) = \sum_H \mathbf{f_4}(DF,CD,H) \times \mathbf{f_5}(H) = \sum_H \left\{ \begin{bmatrix} 0.075 & 0.495 \\ 0.425 & 0.005 \end{bmatrix} \begin{bmatrix} 0.005 & 0.050 \\ 0.495 & 0.450 \end{bmatrix} \right\} = \begin{bmatrix} 0.080 & 0.545 \\ 0.920 & 0.455 \end{bmatrix},$$

$$\mathbf{f_2}(DF) \times \mathbf{f_3}(DF) \times \mathbf{f_6}(DF,CD) = \begin{bmatrix} 0.0096 & 0.0654 \\ 0.0345 & 0.0171 \end{bmatrix},$$

$$\mathbf{f_7}(CD) = \sum_{DF} \mathbf{f_2}(DF) \times \mathbf{f_3}(DF) \times \mathbf{f_6}(DF,CD) = \sum_{DF} \begin{bmatrix} 0.0096 & 0.0654 \\ 0.0345 & 0.0171 \end{bmatrix} = \begin{bmatrix} 0.0441 & 0.0825 \end{bmatrix}.$$

And finally we get:
$\mathbf{P}(CD|\neg a, gf) = \alpha \mathbf{f_1}(CD) \times \mathbf{f_7}(CD) \approx \alpha \begin{bmatrix} 0.0265 & 0.0330 \end{bmatrix}$,
$\alpha = \frac{1}{0.0265+0.0330} \approx 16.8067$,
$\mathbf{P}(CD|\neg a, gf) \approx 16.8067 \cdot \begin{bmatrix} 0.0265 & 0.0330 \end{bmatrix} \approx \begin{bmatrix} 0.4454 & 0.5546 \end{bmatrix}$.

d. Besides the last calculation to complete the normalization that both methods require, we have:

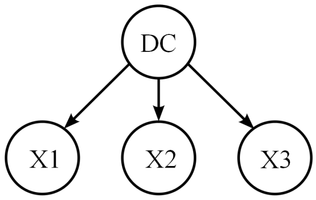
	Enumeration	Variable Elimination
additions	6	6
multiplications	18	16

In this case we do not have a significant improvement using variable elimination. The improvement becomes more significant for larger Bayesian networks.

Problem 8.3:

a. The network is composed of four random variables: The coin that has been drawn DC with domain $< a, b, c >$ and the results of the three flips $X1$, $X2$ and $X3$ with domains $< heads, tails >$.

$P(DC = a)$	$P(DC = b)$	$P(DC = c)$
$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$



DC	$\mathbf{P}(X1 = heads\|DC)$
a	0.2
b	0.6
c	0.8

DC	$\mathbf{P}(X2 = heads\|DC)$
a	0.2
b	0.6
c	0.8

DC	$\mathbf{P}(X3 = heads\|DC)$
a	0.2
b	0.6
c	0.8

b. The required probability is $\mathbf{P}(DC|X1 = heads, X2 = heads, X3 = tails)$.
We first write the full joint probability:
$\mathbf{P}(DC, X1, X2, X3) = \mathbf{P}(X1|DC)\mathbf{P}(X2|DC)\mathbf{P}(X3|DC)\mathbf{P}(DC)$.
The result is directly obtained introducing the evidence and using normalization:

$$\mathbf{P}(DC|X1 = heads, X2 = heads, X3 = tails) =$$
$$= \alpha\mathbf{P}(X1 = heads|DC)\mathbf{P}(X2 = heads|DC)\mathbf{P}(X3 = tails|DC)\mathbf{P}(DC) =$$
$$= \alpha \begin{bmatrix} 0.2 \\ 0.6 \\ 0.8 \end{bmatrix} \times \begin{bmatrix} 0.2 \\ 0.6 \\ 0.8 \end{bmatrix} \times \begin{bmatrix} 0.8 \\ 0.4 \\ 0.2 \end{bmatrix} \times \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \approx \alpha \begin{bmatrix} 0.0107 \\ 0.0480 \\ 0.0427 \end{bmatrix},$$

NB Normalization 'α' is calculated only once

$$\alpha = \frac{1}{(0.0107+0.0480+0.0427)} \approx 9.8619329,$$

$$\mathbf{P}(DC|X1 = heads, X2 = heads, X3 = tails) \approx 9.8619329 \begin{bmatrix} 0.0107 \\ 0.0480 \\ 0.0427 \end{bmatrix} = \begin{bmatrix} 0.1055 \\ 0.4734 \\ 0.4211 \end{bmatrix}.$$

The coin that is most likely to have been drawn is b.