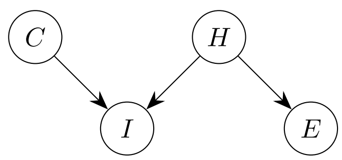
Exercise Session – Solutions

1 Fred's Lisp Dilemma

1. **Modeling Bayesian networks**

   - From the problem statement we get the following graphical representation of the model and its corresponding probability tables.



$p(C)$	C	
	bug	no_bug
	0.4	0.6

$p(H)$	H	
	good	bad
	0.99	0.01

$p(E\|H)$		H	
		good	bad
E	running	0.95	0.1
	not_running	0.05	0.9

I	C	H	$p(I\|C,H)$
responding	bug	good	0.6
not_responding	bug	good	0.4
responding	bug	bad	0.01
not_responding	bug	bad	0.99
responding	no_bug	good	0.9
not_responding	no_bug	good	0.1
responding	no_bug	bad	0.05
not_responding	no_bug	bad	0.95

   - In order to factorize the joint probability distribution we utilize the general formula

$$p(x_1, ..., x_N) = \prod_{i=1}^{N} p(x_i | parents(x_i))$$

     to achieve following factorisation of the model

$$p(C, H, I, E) = p(C)p(H)p(I|H, C)p(E|H).$$

   - When using the full joint, we don't any have information about independence, hence the total number of values needed is $2^4 - 1 = 15$.

   - By utilising the network structure, we can see the independencies between nodes and thus the number of values needed decreases to $2^0 + 2^0 + 2^2 + 2^1 = 8$.

2. Let $I = 0$ mean that the interpreter has stopped running and $C = 0$ that there is a bug in

the code. Then

$$
\begin{aligned}
p(C=0|I=0) &= \frac{p(C=0, I=0)}{p(I=0)} \\
&= \frac{\sum_{H,E} p(C=0, I=0, H, E)}{\sum_{H,E,C} p(I=0, H, E, C)} \\
&= \frac{\sum_{H,E} p(C=0)p(H)p(E|H)p(I=0|C=0,H)}{\sum_{H,E,C} p(C)p(H)p(E|H)p(I=0|C,H)}
\end{aligned}
$$

and by *pushing* the summation over E inside so that it sums to 1 (for a given H, which is fixed by the remaining outer summation)

$$
\begin{aligned}
&= \frac{\sum_{H} p(C=0)p(H)p(I=0|C=0,H)\sum_E p(E|H)}{\sum_{H,C} p(C)p(H)p(I=0|C,H)\sum_E p(E|H)} \\
&= \frac{p(C=0)\sum_H p(I=0|C=0,H)p(H)}{p(C=0)\sum_H p(I=0|C=0,H)p(H) + p(C=1)\sum_H p(I=0|C=1,H)p(H)}
\end{aligned}
$$

numerator and left side of denominator are equal, so we divide them out (for numerical stability)

$$
\begin{aligned}
&= \frac{1}{1+\frac{p(C=1)\sum_H p(I=0|C=1,H)p(H)}{p(C=0)\sum_H p(I=0|C=0,H)p(H)}} \\
&= \frac{1}{1+\frac{0.4(0.4\cdot 0.99+0.99\cdot 0.01)}{0.6(0.1\cdot 0.99+0.95\cdot 0.01)}} \\
&\approx 0.7138
\end{aligned}
$$

Switch.

3. **Independence in Bayesian networks**
   - Collider I is not in the conditioning set, it blocks the paths between C and $\{H, E\}$:

     $$C \perp\!\!\!\perp H \text{ and } C \perp\!\!\!\perp E$$

   - Non-collider in conditioning set H blocks bath between I and E. Path between C and E remains blocked of just H is given. Path between H and C remains blocked if just E is given:

     $$E \perp\!\!\!\perp I|H \text{ and } E \perp\!\!\!\perp C|H \text{ and } C \perp\!\!\!\perp H|E$$

2 Independence in Bayesian Networks

- **Independencies.** Check whether all paths are blocked to verify these!
  - $t \perp\!\!\!\perp s|d$ is false
  - $l \perp\!\!\!\perp b|s$ is true
  - $a \perp\!\!\!\perp s|l$ is true
  - $a \perp\!\!\!\perp s|l, d$ is false

- **Calculating probabilities.** The general idea is to start from the full factorisation and then push summations inside to break the computation up in smaller parts.

$$\begin{aligned} p(d) &= \sum_{x,e,t,l,b,a,s} p(x,d,e,t,l,b,a,s) \\ &= \sum_{x,e,t,l,b,a,s} p(a)p(s)p(t|a)p(l|s)p(b|s)p(x|e)p(d|e,b)p(e|t,l) \end{aligned}$$

push sum over x to the back and notice that $p(s=0) = p(s=1) = 1/2$

$$\begin{aligned} &= \frac{1}{2} \sum_{a,t,l,s,b} p(a)p(t|a)p(l|s)p(b|s)p(d|e,b)p(e|t,l) \sum_x p(x|e) \\ &= \frac{1}{2} \sum_{a,t,l,s,b} p(a)p(t|a)p(l|s)p(b|s)p(d|e,b)p(e|t,l) \qquad (1) \end{aligned}$$

The rest of this is computed in parts. Let's start by pushing the sum over a to the back, noticing that

$$\sum_a p(a)p(t|a) = \sum_a p(t,a) = p(t) \qquad (2)$$

and computing $p(t)$

$$\begin{aligned} p(t=1) &= p(t=1|a=1)p(a=1) + p(t=1|a=0)p(a=0) \\ &= 0.05 \cdot 0.01 + 0.01 \cdot 0.99 \\ &= 0.0104 \\ p(t=0) &= 1 - 0.0104 \end{aligned}$$

From (1) and (2) and pushing the sum over t to the back we get

$$p(d) = \frac{1}{2} \sum_{l,b,e,s} p(l|s)p(b|s)p(d|b,e) \sum_t p(e|t,l)p(t). \qquad (3)$$

Similar to what we did for a, we now find

$$\begin{aligned} \sum_t p(e|t,l)p(t) &= \sum_t \frac{p(e,t,l)p(t)}{p(e,l)} \\ &= \sum_t \frac{p(e,t,l)p(t)}{p(t)p(l)} \text{ because } t \perp\!\!\!\perp l \\ &= \sum_t \frac{p(e,t,l)}{p(l)} \\ &= \frac{p(e,l)}{p(l)} = p(e|l) \qquad (4) \end{aligned}$$

which we can also compute:

$$\begin{aligned} p(e=1|l=1) &= p(e=1|l=1,t=1)p(t=1) + p(e=1|l=1,t=0)p(t=0) \\ &= 1 \cdot 0.0104 + 1 \cdot 0.9896 \\ &= 1 \\ p(e=1|l=0) &= p(e=1|l=0,t=1)p(t=1) + p(e=1|l=0,t=0)p(t=0) \\ &= 1 \cdot 0.0104 + 0 \cdot 0.9896 \\ &= 0.0104 \end{aligned}$$

Combining (3) and (4) and now pushing the sum over l to the back we arrive at

$$p(d) = \frac{1}{2} \sum_{b,e,s} p(b|s)p(d|b,e) \sum_{l} p(e|l)p(l|s). \tag{5}$$

Again, lets see the last term

$$\begin{aligned}
\sum_{l} p(e|l)p(l|s) &= \sum_{l} p(e|l,s)p(l|s) \text{ because } e \perp\!\!\!\perp s|l \\
&= \sum_{l} \frac{p(e,l,s)}{p(s)} \\
&= \frac{p(e,s)}{p(s)} = p(e|s) \qquad (6)
\end{aligned}$$

and compute its probabilities

$$\begin{aligned}
p(e=1|s=1) &= p(e=1|l=1)p(l=1|s=1) + p(e=1|l=0)p(l=0|s=1) \\
&= 1 \cdot 0.1 + 0.0104 \cdot 0.9 \\
&= 0.10936 \\
p(e=1|s=0) &= p(e=1|l=1)p(l=1|s=0) + p(e=1|l=0)p(l=0|s=0) \\
&= 1 \cdot 0.01 + 0.0104 \cdot 0.99 \\
&= 0.020296
\end{aligned}$$

Using (5) and (6) and pushing in e we then arrive at

$$p(d) = \frac{1}{2} \sum_{b,s} p(b|s) \sum_{e} p(d|b,e)p(e|s) \tag{7}$$

for which the last sum will be

$$\begin{aligned}
\sum_{e} p(d|b,e)p(e|s) &= \sum_{e} p(d|b,e,s)p(e|s,b) \text{ because } d \perp\!\!\!\perp s|b,e \text{ and } e \perp\!\!\!\perp b|s \\
&= \sum_{e} \frac{p(d,b,e,s)}{p(b,s)} \\
&= \frac{p(d,b,s)}{p(b,s)} = p(d|b,s), \qquad (8)
\end{aligned}$$

of which we again compute the values.

$$\begin{aligned}
p(d=1|b=1,s=1) &= 0.01093 \\
p(d=1|b=1,s=1) &= 0.80203 \\
p(d=1|b=1,s=1) &= 0.16562 \\
p(d=1|b=1,s=1) &= 0.11217
\end{aligned}$$

Almost there! Plugging (8) into (7) and pushing in b gives

$$p(d) = \frac{1}{2} \sum_{s} \sum_{b} p(d|b,s)p(b|s) \tag{9}$$

for which – you should be getting the hang of it now so details are left out –

$$\sum_b p(d|b,s)p(b|s) = p(d|s) \tag{10}$$

which can be computed, yielding the solution to the two latest asked probability distributions in the process.

$$p(d=1|s=1) = 0.55281$$
$$p(d=1|s=0) = 0.31913$$

Finally, plugging (10) into (9) gives

$$p(d) = \frac{1}{2}\sum_s p(d|s) \tag{11}$$

which is easily computed to give the final result:

$$p(d=1) = \frac{1}{2}(0.55281 + 0.31913) = 0.43597$$
$$p(d=0) = 1 - p(d=1) = 0.56403$$

3 Independence of Random Variables

1. True. They are **unconditionally** independent as we can notice that the $p(A,S)$ is split up in $p(A)p(S)$.
2. False. We see that C is a collider, which means that if it's given it will explicitly make A and S dependent.
3. According to the question, working in the building industry increases the exposure to asbestos as well as increases the likelihood of smoking, hence the new model would be:

$$\begin{aligned} p(A,C,S,B) &= p(A,C,S|B)P(B) \\ &= p(B)p(S|B)p(A|B)p(C|A,S,B) \text{ using the given factorisation} \\ &= p(B)p(S|B)p(A|B)p(C|A,S) \text{ because } C \perp\!\!\!\perp B|A,S \end{aligned}$$

4. The belief network: