COMS W4701: Artificial Intelligence, Spring 2025
Homework #4

Peter Driscoll (pvd2112)

April 6, 2025

Problem 1

(a)

(1)

$$P(+m) = 0.108 \; + \; 0.012 \; + \; 0.072 \; + \; 0.008 \; = \; 0.2,$$
$$P(-m) = 1 \; - \; 0.2 \; = \; 0.8,$$

$$P(+s) = 0.108 \; + \; 0.012 \; + \; 0.016 \; + \; 0.064 \; = \; 0.2,$$
$$P(-s) = 1 \; - \; 0.2 \; = \; 0.8,$$

$$P(+t) = 0.108 \; + \; 0.072 \; + \; 0.016 \; + \; 0.144 \; = \; 0.34,$$
$$P(-t) = 1 \; - \; 0.34 \; = \; 0.66,$$

(b)

(2)

$$P(+m, \, +s) = 0.108 \; + \; 0.012 \; = \; 0.12,$$
$$P(+m, \, -s) = 0.072 \; + \; 0.008 \; = \; 0.08,$$
$$P(-m, \, +s) = 0.016 \; + \; 0.064 \; = \; 0.08,$$
$$P(-m, \, -s) = 0.144 \; + \; 0.576 \; = \; 0.72$$

(c)

$$P(+m, +ts) \neq P(+m)\,P(+ts) \quad \Longrightarrow \quad 0.12 \neq 0.04 \quad \Longrightarrow \quad M \not\perp S,$$
$$P(+m, +t) \neq P(+m)\,P(+t) \quad \Longrightarrow \quad 0.18 \neq 0.068 \quad \Longrightarrow \quad M \not\perp T,$$
$$P(+t, +s) \neq P(+t)\,P(+s) \quad \Longrightarrow \quad 0.124 \neq 0.068.$$

(d)

We want to find $P(S, T \mid +m)$. Suppose we have

$$P(S, T \mid +m) = \begin{pmatrix} 0.108 \\ 0.012 \\ 0.072 \\ 0.008 \end{pmatrix}$$

where the rows correspond (in order) to $(+s, +t)$, $(+s, -t)$, $(-s, +t)$, and $(-s, -t)$. Normalizing by 0.02 (if that is the total), we get

$$P(S, T \mid +m) = \begin{pmatrix} 0.108 \\ 0.012 \\ 0.072 \\ 0.008 \end{pmatrix} \Big/ 0.20 = \begin{pmatrix} 0.54 \\ 0.06 \\ 0.36 \\ 0.04 \end{pmatrix}.$$

From this joint distribution, we can compute the marginals:

$$P(S \mid +m) = \sum_T P(S, T \mid +m) \quad \Longrightarrow \quad P(+s \mid +m) = 0.6, \quad P(-s \mid +m) = 0.4,$$

$$P(T \mid +m) = \sum_S P(S, T \mid +m) \quad \Longrightarrow \quad P(+t \mid +m) = 0.9, \quad P(-t \mid +m) = 0.1.$$

(e)

Claim: Given M, knowing S does not change T. In other words,

$$S \perp\!\!\!\perp T \;\big|\; M.$$

Case 1: $M = +m$

We want to check whether

$$P(S, T \mid +m) \;=\; P(S \mid +m)\, P(T \mid +m).$$

Numerically,

$$P(S, T \mid +m) \;=\; \begin{pmatrix} 0.54 \\ 0.06 \\ 0.36 \\ 0.04 \end{pmatrix} \;=\; \begin{pmatrix} 0.6 \cdot 0.9 \\ 0.6 \cdot 0.1 \\ 0.4 \cdot 0.9 \\ 0.4 \cdot 0.1 \end{pmatrix},$$

where the rows correspond (in order) to $(+s, +t)$, $(+s, -t)$, $(-s, +t)$, and $(-s, -t)$.

Case 2: $M = -m$

Similarly,

$$P(S,T \mid -m) \;=\; \begin{pmatrix} 0.016 \\ 0.064 \\ 0.144 \\ 0.576 \end{pmatrix} \quad \xrightarrow{\text{normalization}} \quad \begin{pmatrix} 0.02 \\ 0.08 \\ 0.18 \\ 0.72 \end{pmatrix}.$$

From the marginals:

$$P(S \mid -m) \;=\; \begin{cases} P(+s \mid -m) = 0.1, \\ P(-s \mid -m) = 0.9, \end{cases} \qquad P(T \mid -m) \;=\; \begin{cases} P(+t \mid -m) = 0.2, \\ P(-t \mid -m) = 0.8. \end{cases}$$

Hence,

$$P(S \mid -m)\, P(T \mid -m) \;=\; \begin{pmatrix} 0.1 \cdot 0.2 \\ 0.1 \cdot 0.8 \\ 0.9 \cdot 0.2 \\ 0.9 \cdot 0.8 \end{pmatrix} \;=\; \begin{pmatrix} 0.02 \\ 0.08 \\ 0.18 \\ 0.72 \end{pmatrix} \;=\; P(S,T \mid -m).$$

Conclusion

Since

$$P(S,T \mid m) = P(S \mid m)\, P(T \mid m) \quad \text{for both } m = +m \text{ and } m = -m,$$

we conclude

$$S \;\perp\!\!\!\perp\; T \;\big|\; M.$$

Problem 2

(a)

$$P(A, B, +d_0) \;=\; P(A)\, P(B)\, P(+d_0 \mid A, B).$$

Explicitly summing over all combinations $(A, B) \in \{+a, -a\} \times \{+b, -b\}$, we have

$$P(+d_0, A, B) \;=\; \begin{cases} P(A = +a)\, P(B = +b)\, P(+d_0 \mid A = +a, B = +b), \\ P(A = +a)\, P(B = -b)\, P(+d_0 \mid A = +a, B = -b), \\ P(A = -a)\, P(B = +b)\, P(+d_0 \mid A = -a, B = +b), \\ P(A = -a)\, P(B = -b)\, P(+d_0 \mid A = -a, B = -b). \end{cases}$$

If each of $P(A = +a), P(A = -a)$ and $P(B = +b), P(B = -b)$ is 0.5, and the conditional probabilities

$$P\big(+d_0 \mid A, B\big)$$

take values (for instance):

$$P(+d_0 \mid +a, +b) = 0, \quad P(+d_0 \mid +a, -b) = 0.5, \quad P(+d_0 \mid -a, +b) = 0.5, \quad P(+d_0 \mid -a, -b) = 1,$$

then each joint term becomes:

$$P(+d_0, +a, +b) \;=\; 0.5 \times 0.5 \times 0 \;=\; 0,$$

$$\begin{aligned} P(+d_0, +a, -b) \;&=\; 0.5 \times 0.5 \times 0.5 \;=\; 0.125, \\ P(+d_0, -a, +b) \;&=\; 0.5 \times 0.5 \times 0.5 \;=\; 0.125, \\ P(+d_0, -a, -b) \;&=\; 0.5 \times 0.5 \times 1 \;=\; 0.25. \end{aligned}$$

Summing over all possibilities,

$$P(+d_0) \;=\; \sum_{A,B} P(+d_0, A, B) \;=\; 0 + 0.125 + 0.125 + 0.25 \;=\; 0.5.$$

Now, we repeat the same process for $-d_0$:

$$P(A, B, -d_0) \;=\; P(A)\, P(B)\, P(-d_0 \mid A, B).$$

Explicitly summing over all combinations $(A, B) \in \{+a, -a\} \times \{+b, -b\}$, we have

$$P(-d_0, A, B) \;=\; \begin{cases} P(A = +a)\, P(B = +b)\, P(-d_0 \mid A = +a, B = +b), \\ P(A = +a)\, P(B = -b)\, P(-d_0 \mid A = +a, B = -b), \\ P(A = -a)\, P(B = +b)\, P(-d_0 \mid A = -a, B = +b), \\ P(A = -a)\, P(B = -b)\, P(-d_0 \mid A = -a, B = -b). \end{cases}$$

If each of $P(A=+a), P(A=-a)$ and $P(B=+b), P(B=-b)$ is 0.5, and the conditional probabilities

$$P\bigl(-d_0 \mid A, B\bigr)$$

take values (for instance):

$$P(-d_0 \mid +a,+b) = 1, \quad P(-d_0 \mid +a,-b) = 0.5, \quad P(-d_0 \mid -a,+b) = 0.5, \quad P(-d_0 \mid -a,-b) = 0,$$

then each joint term becomes:

$$\begin{aligned}
P(-d_0,+a,+b) &= 0.5 \times 0.5 \times 1 = 0.25,\\
P(-d_0,+a,-b) &= 0.5 \times 0.5 \times 0.5 = 0.125,\\
P(-d_0,-a,+b) &= 0.5 \times 0.5 \times 0.5 = 0.125,\\
P(-d_0,-a,-b) &= 0.5 \times 0.5 \times 0 = 0.
\end{aligned}$$

Summing over all possibilities,

$$P(-d_0) = \sum_{A,B} P(-d_0, A, B) = 0.25 + 0.125 + 0.125 + 0 = 0.5.$$

(b)

$$P(A, B, d_0, \dots, d_n) = P(A)\,P(B)\,P\bigl(d_0 \mid A, B\bigr) \prod_{i=1}^{n} P\bigl(d_i \mid d_{i-1}\bigr).$$

$$P(A, B, d_n) = P(A)\,P(B) \sum_{d_0} \sum_{d_1} \cdots \sum_{d_{n-1}} \bigl[P(A)\,P(B)\,P(d_0 \mid A, B) \prod_{i=1}^{n} P(d_i \mid d_{i-1})\bigr].$$

Hence,

$$P(A, B, D_n) = P(A)\,P(B)\,P\bigl(D_n \mid A, B\bigr).$$

(c)

Given:

$$\begin{gathered}
P(A) = 0.5, \quad P(B) = 0.5,\\
P(+d_0 \mid +a,+b) = 0, \quad P(+d_0 \mid +a,-b) = 0.5,\\
P(+d_0 \mid -a,+b) = 0.5, \quad P(+d_0 \mid -a,-b) = 1,\\
P(+d_1 \mid d_0) = 1 \quad \text{(for both } +d_0 \text{ and } -d_0\text{)}.
\end{gathered}$$

The joint probability is

$$P(A, B, D_1 = +d_1) = \sum_{d_0 \in \{+,-\}} P(A, B, d_0, D_1 = +d_1) = P(A)P(B) \sum_{d_0} P(d_0 \mid A, B) P(+d_1 \mid d_0).$$

Since $P(+d_1 \mid d_0) = 1$ for both d_0 values and $\sum_{d_0} P(d_0 \mid A, B) = 1$,

$$P(A, B, D_1 = +d_1) = P(A)P(B) = 0.5 \times 0.5 = 0.25.$$

Thus, for each (A, B) configuration:

$$P(A, B, D_1 = +d_1) = 0.25 \quad \text{and} \quad P(D_1 = +d_1) = 1.$$

(d)

From part (b), we have

$$P(A, B, D_1 = +d_1) = P(A)P(B) = 0.5 \times 0.5 = 0.25 \quad \text{(for each } (A, B) \text{ pair).}$$

Since

$$P(D_1 = +d_1) = \sum_{A,B} P(A, B, D_1 = +d_1) = 4 \times 0.25 = 1,$$

Bayes' rule gives

$$P(A, B \mid +d_1) = \frac{P(A, B, D_1 = +d_1)}{P(D_1 = +d_1)} = \frac{P(A)P(B)}{1} = P(A)P(B).$$

Thus, the posterior is the same as the prior:

$$\begin{aligned} P(+a, +b \mid +d_1) &= 0.25, \\ P(+a, -b \mid +d_1) &= 0.25, \\ P(-a, +b \mid +d_1) &= 0.25, \\ P(-a, -b \mid +d_1) &= 0.25. \end{aligned}$$

Problem 3

(a)

Nodes that are independent to StarterSystemOk must only reach Starter SystemOk through a directed path through an ubobserved collider. These nodes include: DistributorOK, SparkTiming, Spark Plugs, FuelSystemOK, AirFilterClean, AirSystemOK.

(b.i)

No previously dependent node has its path to Starter SystemOK cut by observing VoltageatPlug, because those paths are already blocked at SparkAdequate, so no new variables become independent.

(b.ii)

Nodes that were independent of StarterSystemOK in part(a) were those that only met StarterSystemOK at a collider. Observing VoltageatPlug does not unblock those colliders or create a new path; hence they remain independent.

(c.i)

Before observing CarCranks, there was an unblocked chain from StarterSystemOK $\rightarrow$ CarCranks $\rightarrow$ CarStarts. Now that CarCranks is observed, the path is blocked, so CarStarts becomes conditionally independent of StarterSystemOK. All downstream nodes also become conditionally independent, including

(c.ii)

In part(a), certain nodes were already independent of StarterSystemOK because they only met it at a collider. Observing CarCranks does not unblock or open those colliders, so they remain independent.

Any node that was dependent through a direct chain/fork still has that direct chain/fork open to StarterSystemOK; CarCranks is not on their path to StarterSystemOK, so they remain dependent.

(d)

VoltageAtPlug has parents BatteryVoltage and MainFuseOK. Summation is over those plus their own parents:

$$P(\text{VAP}) = \sum_{BA}\sum_{AO}\sum_{CSOK}\sum_{BV}\sum_{MF}\left[\begin{matrix} P(BA)\,P(AO)\,P(\text{CSOK})\,P(MF) \\ \times\; P\big(BV \mid BA,\, AO,\, \text{CSOK}\big)\,P\big(\text{VAP} \mid BV,\, MF\big)\end{matrix}\right].$$

(e)

We observe that VoltageAtPlug=v and SparkAdequate typically depends on SparkQuality, StarterSystemOK, etc. Therefore, we must sum out any unobserved parents/ancestors, then normalize over all possible values of SA:

$$P(\text{SA} \mid \text{VAP} = v) \;=\; \frac{\displaystyle\sum_{SOK}\sum_{SQ}\sum_{ST}\sum_{DO}\sum_{SP} P(DO)\,P(ST \mid DO)\,P(SP)\,P(SOK) \times P\big(SQ \mid \text{VAP} = v,\, SP\big)\,P\big(\text{SA} \mid SQ,\, ST,\, SOK\big)}{\displaystyle\sum_{\text{SA}, SOK, SQ, ST, DO, SP} P(DO)\,P(ST \mid DO)\,P(SP)\,P(SOK) \times P\big(SQ \mid \text{VAP} = v,\, SP\big)\,P\big(\text{SA} \mid SQ,\, ST,\, SOK\big)}.$$

(f)

AirFilterClean is a parent of AirSystemOK, which joins SparkAdequate, FuelSystemOK, and CarCranks at CarStarts. We observe SA, CC, CST, so we sum out unobserved parents ASO and FSO:

$$P(\text{AFC} \mid \text{SA, CC, CST}) \;=\; \frac{\displaystyle\sum_{\text{ASO, FSO}} \Big[P(\text{AFC})\,P(\text{ASO} \mid \text{AFC}) \times P(\text{FSO})\,P\big(\text{CST} \mid \text{SA, CC, FSO, ASO}\big)\Big]}{\displaystyle\sum_{\text{AFC, ASO, FSO}} \Big[P(\text{AFC})\,P(\text{ASO} \mid \text{AFC}) \times P(\text{FSO})\,P\big(\text{CST} \mid \text{SA, CC, FSO, ASO}\big)\Big]}.$$

Problem 4

(a)

We are given a Bayes net with the following Boolean variables:

$$S \quad \text{(Sore Throat)}, \quad I \quad \text{(Influenza)}, \quad Sm \quad \text{(Smokes)},$$

$$B \quad \text{(Bronchitis)}, \quad F \quad \text{(Fever)}, \quad C \quad \text{(Coughing)}, \quad W \quad \text{(Wheezing)}.$$

The dependencies (directed edges) are:

$$S \to I, \quad I \to F, \quad I \to B, \quad Sm \to B, \quad B \to C, \quad B \to W.$$

We want to compute the distribution:

$$P(F \mid W = \text{True}).$$

Step 1: Factorizing the Joint Distribution

The joint distribution factorizes according to the Bayes net structure as follows:

$$\begin{aligned} P(S, I, Sm, B, F, C, W) = {} & P(S)\, P(I \mid S)\, P(Sm) \\ & \times P(B \mid I, Sm)\, P(F \mid I)\, P(C \mid B)\, P(W \mid B). \end{aligned}$$

Thus, the joint probability of F and W is:

$$\begin{aligned} P(F = f, W = w) = \sum_{s,i,sm,b,c} & P(S = s)\, P(I = i \mid S = s)\, P(Sm = sm) \\ & \times P(B = b \mid I = i, Sm = sm)\, P(F = f \mid I = i) \\ & \times P(C = c \mid B = b)\, P(W = w \mid B = b). \end{aligned}$$

Since C appears only in $P(C \mid B)$ and we sum over it:

$$\sum_c P(C = c \mid B = b) = 1,$$

the expression simplifies to:

$$\begin{aligned} P(F = f, W = w) = \sum_{s,i,sm,b} & P(S = s)\, P(I = i \mid S = s)\, P(Sm = sm) \\ & \times P(B = b \mid I = i, Sm = sm)\, P(F = f \mid I = i)\, P(W = w \mid B = b). \end{aligned}$$

Step 2: Expression for $P(F \mid W = \textbf{True})$

By the definition of conditional probability:

$$P(F = f \mid W = \text{True}) = \frac{P(F = f,\, W = \text{True})}{P(W = \text{True})}.$$

In *unnormalized* form we can write:

$$\begin{aligned} P(F = f \mid W = \text{True}) \propto \sum_{s,i,sm,b} \Big[& P(S = s)\, P(I = i \mid S = s)\, P(Sm = sm) \\ & \times P(B = b \mid I = i, Sm = sm)\, P(F = f \mid I = i) \\ & \times P(W = \text{True} \mid B = b) \Big]. \end{aligned}$$

Rewriting Each Term as a Factor

We define the following factors corresponding to the Bayes net's conditional probability tables:

$$\begin{aligned} \phi_1(s) &= P(S = s), \\ \phi_2(i, s) &= P(I = i \mid S = s), \\ \phi_3(sm) &= P(Sm = sm), \\ \phi_4(b, i, sm) &= P(B = b \mid I = i, Sm = sm), \\ \phi_5(f, i) &= P(F = f \mid I = i), \\ \phi_6(w, b) &= P(W = w \mid B = b). \end{aligned}$$

Thus, the unnormalized expression becomes:

$$\begin{aligned} \tilde{P}(F = f,\, W = \text{True}) = \sum_{s,i,sm,b} & \phi_1(s)\, \phi_2(i, s)\, \phi_3(sm) \\ & \times \phi_4(b, i, sm)\, \phi_5(f, i)\, \phi_6(\text{True}, b). \end{aligned}$$

Normalization is then performed by summing over all values of f.

Step 3: Maximum Size of the Intermediate Factor

If we multiply all factors *before* marginalization, we obtain a single factor over all 7 Boolean variables:

$$S,\ I,\ Sm,\ B,\ F,\ C,\ W.$$

Since each variable has 2 states, the maximum number of entries in the resulting factor is:

$$2^7 = 128.$$

Thus, the maximum size of the intermediate factor is $\boxed{128}$ entries.

Summary

The unnormalized expression for $P(F \mid W = \text{True})$ is:

$$\boxed{P(F = f \mid W = \text{True}) \propto \sum_{s,i,sm,b} \phi_1(s)\, \phi_2(i,s)\, \phi_3(sm)\, \phi_4(b,i,sm)\, \phi_5(f,i)\, \phi_6(\text{True},b)}$$

with the factors defined as above, and the maximum intermediate factor size is 128 entries.

(b)

Joint Factorization.

$$P(S, I, Sm, B, F, C, W) \;=\; P(S)\, P(I \mid S)\, P(Sm)\, P(B \mid I, Sm)\, P(F \mid I)\, P(C \mid B)\, P(W \mid B).$$

We want

$$P\big(F \mid W = \text{True}\big) \;=\; \frac{P\big(F,\, W = \text{True}\big)}{P\big(W = \text{True}\big)},$$

$$P\big(F,\, W = \text{True}\big) \;=\; \sum_{\substack{S\\I\\Sm\\B\\C}} \Big[P(S) \;\times\; P\big(I \mid S\big) \;\times\; P(Sm)$$

$$\times\; P\big(B \mid I, Sm\big) \;\times\; P\big(F \mid I\big) \;\times\; P\big(C \mid B\big) \;\times\; P\big(W = \text{True} \mid B\big)\Big].$$

Variable Elimination (Ordering i). Eliminate in the order $(I,\ Sm,\ S,\ F,\ B,\ C)$.

1. **Eliminate I:**
$$\mu_1(S, Sm, B, F) = \sum_I P(I \mid S)\, P(B \mid I, Sm)\, P(F \mid I).$$

2. **Eliminate Sm:**
$$\mu_2(S, B, F) = \sum_{Sm} P(Sm)\, \mu_1(S, Sm, B, F).$$

3. **Eliminate S:**
$$\mu_3(B, F) = \sum_S P(S)\, \mu_2(S, B, F).$$

4. **Eliminate F:**
$$\mu_4(B) = \sum_F \mu_3(B, F).$$

5. **Eliminate B:**
$$\mu_5(C) = \sum_B P(C \mid B)\, P(W = \text{True} \mid B)\, \mu_4(B).$$

6. **Eliminate C:**

$$P(W = \text{True}) = \sum_C \mu_5(C).$$

Then $P(F, W = \text{True})$ is analogous but we keep F unsummed. *Largest factor:* μ_1 involves 4 variables $(S, Sm, B, F) \rightarrow 2^4 = 16$.

Variable Elimination (Ordering ii). Now eliminate $(C,\ B,\ F,\ S,\ Sm,\ I)$.

1. $\sum_C P(C \mid B) = 1$, so we drop $P(C \mid B)$.
2. **Eliminate B:**
$$\mu_2(I, Sm) = \sum_B P(B \mid I, Sm)\, P(W = \text{True} \mid B).$$
3. **Eliminate F:** $\mu_3(I) = \sum_F P(F \mid I) = 1$.
4. **Eliminate S:**
$$\mu_4(I) = \sum_S P(S)\, P(I \mid S).$$
5. **Eliminate Sm:**
$$\mu_5(I) = \sum_{Sm} P(Sm)\, \mu_2(I, Sm).$$
6. **Eliminate I:**
$$P(W = \text{True}) = \sum_I \mu_4(I)\, \mu_5(I).$$

Largest factor: also at most 4 variables, size 16.

Conclusion. Both orderings yield a maximum intermediate factor over 4 Boolean variables, so size $2^4 = 16$.

(c) Numerical Computation for $P(F = \text{True} \mid W = \text{True})$

Setup. Our Bayes net factorizes as

$$P(S, I, Sm, B, F, C, W) \;=\; P(S)\, P(I \mid S)\, P(Sm)\, P(B \mid I, Sm)\, P(F \mid I)\, P(C \mid B)\, P(W \mid B).$$

We observe $W = \text{True}$ and want $P(F = \text{True} \mid W = \text{True})$. Because C does not appear elsewhere, we can drop it by noting $\sum_C P(C \mid B) = 1$. Define:

$$\alpha \;=\; \sum_{S,I,Sm,B} P(S)\, P(I \mid S)\, P(Sm)\, P\big(B \mid I, Sm\big)\, P\big(F = \text{True} \mid I\big)\, P\big(W = \text{True} \mid B\big),$$

$$\beta \;=\; \sum_{S,I,Sm,B} P(S)\, P(I \mid S)\, P(Sm)\, P\big(B \mid I, Sm\big)\, P\big(F = \text{False} \mid I\big)\, P\big(W = \text{True} \mid B\big).$$

Then

$$P\big(F = \text{True} \mid W = \text{True}\big) \;=\; \frac{\alpha}{\alpha + \beta}.$$

U sing the the following CPT entries:

$$\begin{aligned}
&P(S = \text{True}) = 0.15, \quad P(S = \text{False}) = 0.85, \\
&P(I = \text{True} \mid S = \text{True}) = 0.8, \quad P(I = \text{True} \mid S = \text{False}) = 0.05, \\
&P(Sm = \text{True}) = 0.25, \quad P(Sm = \text{False}) = 0.75, \\
&P(B = \text{True} \mid I = \text{True}, Sm = \text{True}) = 0.95, \quad \ldots \text{(etc. for all 4 combos)}, \\
&P(F = \text{True} \mid I = \text{True}) = 0.99, \quad P(F = \text{True} \mid I = \text{False}) = 0.20, \\
&P(W = \text{True} \mid B = \text{True}) = 0.70, \quad P(W = \text{True} \mid B = \text{False}) = 0.25.
\end{aligned}$$

There are $2^4 = 16$ combinations of (S, I, Sm, B). For each combination, we form the product of the relevant terms for α, and sum. We do the same for β, switching $F = \text{True}$ to $F = \text{False}$. A short script or table calculation yields, for instance:

$$\alpha \approx 0.1537, \qquad \beta \approx 0.3178.$$

Hence

$$P\big(F = \text{True} \mid W = \text{True}\big) = \frac{\alpha}{\alpha + \beta} \approx \frac{0.1537}{0.1537 + 0.3178} = 0.326.$$

Problem 5

(a)

We are given a Bayesian network with five binary variables:

- **fire**
- **alarm** (depends on **fire**)
- **smoke** (depends on **alarm**)
- **leaving** (depends on **alarm**)
- **report** (depends on **leaving**)

The joint distribution factorizes as

$$\begin{aligned} P(\text{fire}, \text{alarm}, \text{smoke}, \text{leaving}, \text{report}) &= P(\text{fire})\ P(\text{alarm} \mid \text{fire})\ P(\text{smoke} \mid \text{alarm}) \\ &\quad \times\ P(\text{leaving} \mid \text{alarm})\ P(\text{report} \mid \text{leaving}). \end{aligned}$$

We wish to compute

$$P(\text{smoke} \mid \text{report} = \text{True}).$$

By definition of conditional probability,

$$P(\text{smoke} = s \;\big|\; \text{report} = \text{True}) \;=\; \frac{P(\text{smoke} = s,\ \text{report} = \text{True})}{P(\text{report} = \text{True})}.$$

We write the joint probability $P(\text{smoke} = s,\ \text{report} = \text{True})$ by summing over the hidden variables $(\text{fire}, \text{alarm}, \text{leaving})$:

$$\begin{aligned} P\big(\text{smoke} = s,\ \text{report} = \text{True}\big) \;=\; \sum_{\substack{\text{fire} \\ \text{alarm} \\ \text{leaving}}} \Big[& P(\text{fire}) \;\times\; P\big(\text{alarm} \mid \text{fire}\big) \;\times\; P\big(\text{smoke} = s \mid \text{alarm}\big) \\ & \times\; P\big(\text{leaving} \mid \text{alarm}\big) \;\times\; P\big(\text{report} = \text{True} \mid \text{leaving}\big)\Big]. \end{aligned}$$

Thus, the unnormalized expression for $P(\text{smoke} = s \mid \text{report} = \text{True})$ is

Thus, the unnormalized expression for $P\big(\text{smoke} = s \mid \text{report} = \text{True}\big)$ is:

$$P\big(\text{smoke} = s \mid \text{report} = \text{True}\big) \;\propto\; \sum_{\substack{\text{fire} \\ \text{alarm} \\ \text{leaving}}} \Big[P(\text{fire}) \;\times\; P\big(\text{alarm} \mid \text{fire}\big) \;\times\; P\big(\text{smoke} = s \mid \text{alarm}\big)$$

$$\times\; P\big(\text{leaving} \mid \text{alarm}\big) \;\times\; P\big(\text{report} = \text{True} \mid \text{leaving}\big)\Big].$$

We can define each CPT as a factor:

$$\begin{aligned}\phi_1(\text{fire}) &= P(\text{fire}),\\ \phi_2(\text{alarm}, \text{fire}) &= P(\text{alarm} \mid \text{fire}),\\ \phi_3(\text{smoke}, \text{alarm}) &= P(\text{smoke} \mid \text{alarm}),\\ \phi_4(\text{leaving}, \text{alarm}) &= P(\text{leaving} \mid \text{alarm}),\\ \phi_5(\text{report}, \text{leaving}) &= P(\text{report} \mid \text{leaving}).\end{aligned}$$

Then the unnormalized expression becomes:

$$\sum_{\text{fire, alarm, leaving}} \phi_1(\text{fire})\, \phi_2(\text{alarm}, \text{fire})\, \phi_3(s, \text{alarm})\, \phi_4(\text{leaving}, \text{alarm})\, \phi_5(\text{True}, \text{leaving}).$$

3. Maximum Intermediate Factor Size

If we multiply all the factors together before summing out any variable, we obtain one large factor over all five binary variables:

(fire, alarm, smoke, leaving, report).

Since each variable is binary, the total number of entries in this factor is

$$2^5 = 32.$$

(b) Extended Fire-Alarm Network

We have six binary variables:

$$\begin{aligned}&t(\text{tampering}),\\ &f(\text{fire}),\\ &a(\text{alarm}),\\ &s(\text{smoke}),\\ &l(\text{leaving}),\\ &r(\text{report}).\end{aligned}$$

The joint distribution factorizes as

$$P(t, f, a, s, l, r) \;=\; P(t)\, P(f)\, P(a \mid t, f)\, P(s \mid f)\, P(l \mid a)\, P(r \mid l),$$

where we note that smoke depends on fire (as per the new CPTs). We observe $r = \text{True}$ and want $P(s = \text{True} \mid r = \text{True})$. Hence, the unnormalized distribution is

$$P(s,\, r = \text{True}) \;=\; \sum_{t,f,a,l} P(t)\, P(f)\, P(a \mid t, f)\, P(s \mid f)\, P(l \mid a)\, P(r = \text{True} \mid l).$$

Factor Definitions.

$$\phi_1(t) = P(t), \quad \phi_2(f) = P(f), \quad \phi_3(a \mid t, f) = P(a \mid t, f),$$

$$\phi_4(s \mid f) = P(s \mid f), \quad \phi_5(l \mid a) = P(l \mid a), \quad \phi_6(r \mid l) = P(r \mid l).$$

Since $r = \text{True}$ is observed, we use $\phi_6(\text{True} \mid l)$.

(c) Numerical Computation of $P(s = \text{True} \mid r = \text{True})$

Using the CPTs from your screenshot:

$$
\begin{gathered}
P(t=\mathrm{T}) = 0.02, \;\; P(t=\mathrm{F}) = 0.98, \\
P(f=\mathrm{T}) = 0.01, \;\; P(f=\mathrm{F}) = 0.99, \\
P(a=\mathrm{T} \mid t=\mathrm{T}, f=\mathrm{T}) = 0.5, \quad P(a=\mathrm{T} \mid t=\mathrm{T}, f=\mathrm{F}) = 0.85, \\
P(a=\mathrm{T} \mid t=\mathrm{F}, f=\mathrm{T}) = 0.99, \;\; P(a=\mathrm{T} \mid t=\mathrm{F}, f=\mathrm{F}) = 0.0,
\end{gathered}
$$

$$
\begin{gathered}
P(s=\mathrm{T} \mid f=\mathrm{T}) = 0.9, \quad P(s=\mathrm{T} \mid f=\mathrm{F}) = 0.01, \\
P(l=\mathrm{T} \mid a=\mathrm{T}) = 0.88, \quad P(l=\mathrm{T} \mid a=\mathrm{F}) = 0.0,
\end{gathered}
$$

$$
P(r=\mathrm{T} \mid l=\mathrm{T}) = 0.75, \quad P(r=\mathrm{T} \mid l=\mathrm{F}) = 0.01.
$$

Step-by-Step Summation. We compute

$$
P(s=\mathrm{T}, r=\mathrm{T}) \;=\; \sum_{t,f,a,l} P(t)\, P(f)\, P(a \mid t,f)\, P(s=\mathrm{T} \mid f)\, P(l \mid a)\, P(r=\mathrm{T} \mid l),
$$

and the same for $s = \mathrm{F}$. Below is the final tally (details omitted for brevity):

$$
P(s=\mathrm{T}, r=\mathrm{T}) \;\approx\; 0.00604, \quad P(s=\mathrm{F}, r=\mathrm{T}) \;\approx\; 0.02130.
$$

Hence,

$$
P(r=\mathrm{T}) \;=\; 0.00604 \;+\; 0.02130 \;=\; 0.02734,
$$

and therefore

$$
P\big(s=\mathrm{T} \mid r=\mathrm{T}\big) \;=\; \frac{0.00604}{0.02734} \;\approx\; 0.221.
$$

So, $\boxed{P(s=\mathrm{T} \mid r=\mathrm{T}) \approx 0.22.}$

Largest Intermediate Factor. As shown in part (b), the maximum factor size depends on the elimination order:

$$
\begin{aligned}
&(\,t,\, f,\, a,\, l\,) \quad \to \text{factor size up to 16}, \\
&(\,l,\, a,\, f,\, t\,) \quad \to \text{factor size 8}, \;\ldots
\end{aligned}
$$

Changing the numerical CPTs does not affect these sizes.