Intro to AI Assignment 3 - Probabilistic Reasoning

Rajeev Atla - 2008003072
Dhvani Patel - 21006030
Jasmin Badyal - 2008003131

May 4, 2025

§1 Problem 1

§1.1 Part a

The joint probability for the events A, B, C, D, E is defined by the chain rule as:

$$P(A, B, C, D, E) = P(A) \cdot P(B) \cdot P(C) \cdot P(D|A, B) \cdot P(E|B, C)$$

For the specific case where all events are true (T), the joint probability is:

$$\begin{aligned} P(A = T, B = T, C = T, D = T, E = T) &= P(A = T) \cdot P(B = T) \cdot P(C = T) \\ &\cdot P(D = T|A = T, B = T) \cdot P(E = T|B = T, C = T) \end{aligned}$$

Substituting the provided numerical values:

$$\begin{aligned} P(A = T, B = T, C = T, D = T, E = T) &= 0.2 \cdot 0.5 \cdot 0.8 \cdot 0.1 \cdot 0.3 \\ &= 0.0024 \end{aligned}$$

Thus, the likelihood of the combined event where A, B, C, D and E are all true is $\boxed{0.0024}$.

§1.2 Part b

Now, let's analyze the scenario where all these events are false (F). The joint probability is:

$$\begin{aligned} P(A = F, B = F, C = F, D = F, E = F) &= P(A = F) \cdot P(B = F) \cdot P(C = F) \\ &\cdot P(D = F|A = F, B = F) \cdot P(E = F|B = F, C = F) \end{aligned}$$

Inserting the given probabilities for this scenario:

$$\begin{aligned} P(A = F, B = F, C = F, D = F, E = F) &= 0.8 \cdot 0.5 \cdot 0.2 \cdot 0.1 \cdot 0.8 \\ &= 0.0064 \end{aligned}$$

Consequently, the probability of the joint event where A, B, C, D, and E are all false is $\boxed{0.0064}$.

§1.3 Part c

We aim to determine the conditional probability $P(\neg A|B,C,D,E)$. Employing Bayes' theorem, this can be expressed as proportional to the joint probability $P(\neg A,B,C,D,E)$:

$$P(\neg A|B,C,D,E) \propto P(\neg A,B,C,D,E).$$

Let the normalization factor be α, defined as:

$$\alpha = \frac{1}{P(A,B,C,D,E) + P(\neg A,B,C,D,E)}.$$

We are given the calculation for α:

$$\alpha = \frac{1}{(0.2 \cdot 0.5 \cdot 0.8 \cdot 0.1 \cdot 0.3) + (0.8 \cdot 0.5 \cdot 0.8 \cdot 0.6 \cdot 0.3)}.$$

$$\alpha = \frac{1}{0.0024 + 0.0576} = \frac{1}{0.06} = \frac{50}{3}.$$

Now, we can compute the conditional probability $P(\neg A|B,C,D,E)$:

$$P(\neg A|B,C,D,E) = \alpha \cdot P(\neg A,B,C,D,E).$$

$$P(\neg A|B,C,D,E) = \frac{50}{3} \cdot 0.0576.$$

$$P(\neg A|B,C,D,E) = 0.96.$$

Therefore, the conditional probability $P(\neg A|B,C,D,E)$ is $\boxed{0.96}$.

§2 Problem 2

§2.1 Part a

Our goal is to determine the conditional probability $P(\text{Burglary}|\text{JohnsCalls} = \text{true}, \text{MaryCalls} = \text{true})$. We are given the formulation:

$$P(B|J,M) = \alpha \cdot P(B) \sum_E P(E) \sum_A P(A|B,E) \cdot P(J|A) \cdot P(M|A)$$

where $\alpha = \dfrac{1}{P(J,M)}$.

Following the provided steps and correcting the final calculation:

$$P(B|J,M) = \alpha \cdot \begin{pmatrix} 0.00059224259 \\ 0.0014918576 \end{pmatrix}$$

(where the top element corresponds to $B = T$ and the bottom to $B = F$)

$$\left(\text{and } \alpha = \frac{1}{0.0020853609}\right)$$

$$= \frac{1}{0.0020853609} \cdot \begin{pmatrix} 0.00059224259 \\ 0.0014918576 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{0.00059224259}{0.0020853609} \\ \frac{0.0014918576}{0.0020853609} \end{pmatrix}$$

$$= \begin{pmatrix} 0.284 \\ 0.716 \end{pmatrix}$$

Here, $\boxed{0.284}$ signifies the probability of a burglary occurring given that John and Mary call, while $\boxed{0.716}$ represents the probability of no burglary under the same conditions.

§2.2 Part b

What is the computational cost of determining $P(X_1|X_n = \text{true})$ via enumeration? What is the cost using variable elimination?

§2.2.1 Complexity via Enumeration

To compute $P(X_1|X_n = \text{true})$ by enumeration, we initially assess two binary trees for each state of X_1. Each of these trees possesses a depth of $n - 2$. Consequently, the aggregate computational effort for enumeration amounts to $\boxed{O(2^n)}$.

§2.2.2 Complexity via Variable Elimination

Moving on to variable elimination, the size of the factors will not exceed two variables. For instance, when computing $P(X_1|X_n = \text{true})$:

$$\begin{aligned} P(X_1|X_n = \text{true}) &= \alpha \cdot P(X_1) \cdots \sum_{x_{n-2}} P(x_{n-2}|x_{n-3}) \sum_{x_{n-1}} P(x_{n-1}|x_{n-2}) P(X_n = \text{true}|x_{n-1}) \\ &= \alpha \cdot P(X_1) \cdots \sum_{x_{n-2}} P(x_{n-2}|x_{n-3}) \sum_{x_{n-1}} f_{X_{n-1}}(x_{n-1}, x_{n-2}) f_{X_n}(x_{n-1}) \\ &= \alpha \cdot P(X_1) \cdots \sum_{x_{n-2}} P(x_{n-2}|x_{n-3}) f_{\frac{x_{n-2}}{X_{n-1} \cdot X_n}} \end{aligned}$$

As evident, this mirrors a problem with $n - 1$ variables rather than n. Hence, the computational work remains constant, independent of n, and the overall complexity is $\boxed{O(n)}$.

§3 Problem 3

§4 Problem 4

We can model the system as a hidden Markov model. We can model X_t as a Markov chain with the states $\{A, B, C, D, E, F\}$ and transition matrix:

$$T_{i,j} = P(X_t = j | X_{t-1} = i)$$
$$T = \begin{pmatrix} 0.2 & 0.8 & 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0.8 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 & 0 & 0 \\ 0 & 0 & 0 & 0.2 & 0.8 & 0 \\ 0 & 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

In addition, we have the observation matrices for hot and cold:

$$O_{\text{hot}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$O_{\text{cold}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

We know that the rover starts at state A with probability 1, so $P(X_1 = A) = 1$. The initial state vector is therefore $f_1 = [1, 0, 0, 0, 0, 0]^T$.

§4.1 Part 1

We are being asked to compute the state distribution f_3 given the observations $O_1 = \text{hot}$, $O_2 = \text{cold}$, and $O_3 = \text{cold}$.

$$
\begin{aligned}
f_3 &= \alpha O_{e_3} T f_2 \\
&= \alpha O_{e_3} T \left(\alpha O_{e_2} T f_1\right) \\
&= \alpha O_{\text{cold}} T O_{\text{cold}} T f_1 \\
&= \alpha \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.2 & 0.8 & 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0.8 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 & 0 & 0 \\ 0 & 0 & 0 & 0.2 & 0.8 & 0 \\ 0 & 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
&\cdot \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.2 & 0.8 & 0 & 0 & 0 & 0 \\ 0 & 0.2 & 0.8 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 & 0 & 0 \\ 0 & 0 & 0 & 0.2 & 0.8 & 0 \\ 0 & 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \\
&= \alpha[0, 0.16, 0.64, 0, 0, 0] \\
&= [0, \frac{1}{5}, \frac{4}{5}, 0, 0, 0]
\end{aligned}
$$

§4.2 Part 2

We know that

$$
f_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
$$

We can then compute

$$
\begin{aligned}
b_3 &= T O_{\text{cold}} b_4 \\
&= [0.8, 1, 0.2, 0.8, 1, 0] \\
P(X_2 | \text{hot}_1, \text{cold}_2, \text{cold}_3) &= b_3 \circ f_2 \\
&= [0, 1, 0, 0, 0, 0]
\end{aligned}
$$

§4.4 Part 4

We can simply do

$$\begin{aligned} P(X_4|\text{hot}_1, \text{cold}_2, \text{cold}_3) &= Tf_3 \\ &= [0, 0.04, 0.32, 0.64, 0, 0] \end{aligned}$$

§4.3 Part 3

Using the results of part 4,

$$\begin{aligned} P(\text{hot}_4|\text{hot}_1, \text{cold}_2, \text{cold}_3) &= P(X_4 = A \text{ or } X_4 = D|\text{hot}_1, \text{cold}_2, \text{cold}_3) \\ &= 0 + 0.64 \\ &= 0.64 \end{aligned}$$

§5 Problem 5

§5.1 Part 1

The Bellman equation for this Markov decision process is given by:

$$V^\pi(s) = R(s, \pi(s)) + \sum_{s' \in A,B} T(s, \pi(s), s')V^\pi(s')$$

§5.2 Part 2

§5.2.1 Iteration 1

$$\begin{aligned} V_0^\pi(A) &= 0 \\ V_1^\pi(A) &= R(A, 1) + \sum_{s'} T(A, 1, s')V_0^\pi(s') \\ &= 0 + (1 \cdot 0 + 0 \cdot 0) \\ &= 0 \\ V_0^\pi(B) &= 0 \\ V_1^\pi(B) &= R(B, 1) + \sum_{s'} T(B, 1, s')V_0^\pi(s') \\ &= 5 + (0 \cdot 0 + 1 \cdot 0) \\ &= 5 \end{aligned}$$

§5.2.2 Iteration 2

$$\begin{aligned}
V_2^\pi(A) &= R(A,1) + \sum_{s'} T(A,1,s')V_0^\pi(s') \\
&= 0 + (1 \cdot 0 + 0 \cdot 5) \\
&= 0 \\
V_2^\pi(B) &= R(B,1) + \sum_{s'} T(B,1,s')V_0^\pi(s') \\
&= 5 + (0 \cdot 0 + 1 \cdot 5) \\
&= 10
\end{aligned}$$

§5.3 Part 3

$$\begin{aligned}
\pi_{\text{new}}(A) &= \operatorname*{argmax}_a \left\{ R(A,1) + \sum_{s'} T(A,1,s')V_0^\pi(s'), R(A,2) + \sum_{s'} T(A,2,s')V_0^\pi(s') \right\} \\
&= \operatorname*{argmax}_a \{0 + (1 \cdot 0 + 0 \cdot 5), 0 + (0 \cdot 0 + 1 \cdot 5)\} \\
&= \operatorname*{argmax}_a \{0, 5\} \\
&= 2 \\
\pi_{\text{new}}(B) &= \operatorname*{argmax}_a \left\{ R(B,1) + \sum_{s'} T(B,1,s')V_0^\pi(s'), R(B,2) + \sum_{s'} T(B,2,s')V_0^\pi(s') \right\} \\
&= \operatorname*{argmax}_a \{5 + (0 \cdot 0 + 1 \cdot 5), 5 + (1 \cdot 0 + 0 \cdot 5)\} \\
&= \operatorname*{argmax}_a \{10, 5\} \\
&= 1
\end{aligned}$$