

Homework 1 - Group 03

Planning, Learning and Intelligent Decision Making

Duarte Almeida
95565

Martim Santos
95638

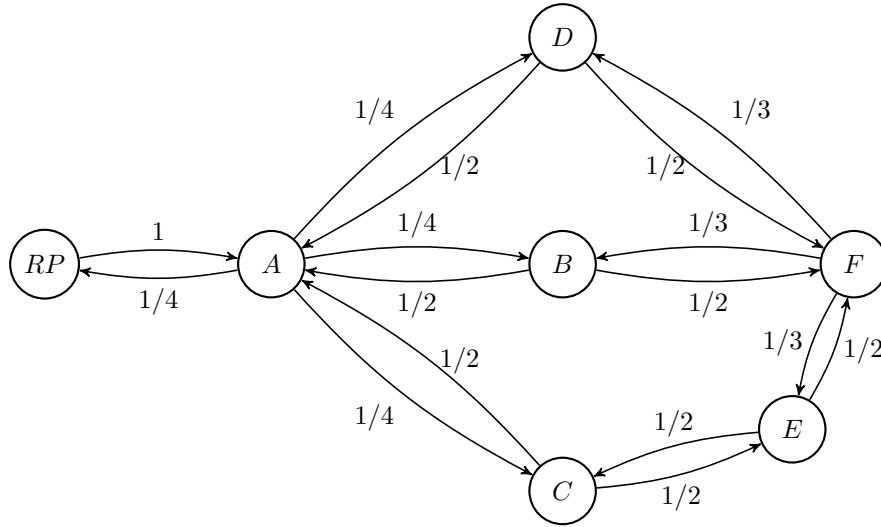
Exercise 1.

(a) The motion of the truck can be described by a Markov Chain $\mathcal{M} = (\mathcal{X}, \mathbf{P})$, where:

$$\mathcal{X} = \{RP, A, B, C, D, E, F\}$$

$$\mathbf{P} = \begin{matrix} & \begin{matrix} RP & A & B & C & D & E & F \end{matrix} \\ \begin{matrix} RP \\ A \\ B \\ C \\ D \\ E \\ F \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \end{bmatrix} \end{matrix}$$

A diagram representing this Markov Chain is:



(b) By the Chapman-Kolmogorov equations, we have, for $x \in \mathcal{X}$:

$$\begin{aligned} \mathbb{P}(x_2 = x \mid x_0 = RP) &= \sum_{y \in \mathcal{X}} \mathbb{P}(x_1 = y \mid x_0 = RP) \mathbb{P}(x_2 = x \mid x_1 = y) \\ &= \mathbb{P}(x_2 = x \mid x_1 = A) = \mathbf{P}(x \mid A) \end{aligned}$$

with the last equality coming from the fact that A is the only possible state for time step 1 if the initial state is RP . Thus, we have that the required probabilities are:

$$[\mathbf{P}][0, :] = \begin{matrix} & \begin{matrix} RP & A & B & C & D & E & F \end{matrix} \\ \begin{matrix} RP \end{matrix} & \begin{bmatrix} 0.25 & 0 & 0.25 & 0.25 & 0.25 & 0 & 0 \end{bmatrix} \end{matrix}$$

- (c) Let $E = \{(x, y) \in \mathcal{X} \times \mathcal{X} : \mathbf{P}(y | x) > 0\}$ denote the set of pairs of states that correspond to sites with roads between them. Let $T_{xy|RP}$ be a random variable that expresses the time spent travelling from site x to site y before returning to the recycling plant. We are thus asked to find:

$$\mathbb{E} \left[\sum_{(x,y) \in E} T_{xy|RP} \right] = \mathbb{E} \left[\sum_{(x,y) \in E} t_{xy} N_{xy|RP} \right] = \sum_{(x,y) \in E} t_{xy} \underbrace{\mathbb{E} [N_{xy|RP}]}_{\eta_{xy|RP}} \quad (1)$$

where t_{xy} denotes the time needed to travel from site x to site y and $N_{xy|x^*}$ is a random variable that expresses the number of times the truck travels from site x to site y before returning to the recycling plant. Let also T_{x^*} denote the first return time to x^* . We then have:

$$\begin{aligned} \eta_{xy|x^*} &= \mathbb{E} \left[\sum_{t=1}^{\infty} \mathbb{I}[x_t = x, x_{t+1} = y, T_{x^*} > t \mid x_0 = x^*] \right] \\ &= \sum_{t=1}^{\infty} \mathbb{P}[x_t = x, x_{t+1} = y, T_{x^*} > t \mid x_0 = x^*] \\ &= \sum_{t=1}^{\infty} \mathbb{P}[x_{t+1} = y \mid x_t = x, T_{x^*} > t, x_0 = x^*] \mathbb{P}[x_t = x, T_{x^*} > t \mid x_0 = x^*] \\ &= \sum_{t=1}^{\infty} \mathbb{P}[x_{t+1} = y \mid x_t = x, x_0 = x^*] \mathbb{P}[x_t = x, T_{x^*} > t \mid x_0 = x^*] \end{aligned} \quad (2)$$

$$\begin{aligned} &= \sum_{t=1}^{\infty} \mathbb{P}[x_{t+1} = y \mid x_t = x] \mathbb{P}[x_t = x, T_{x^*} > t \mid x_0 = x^*] \\ &= \sum_{t=1}^{\infty} \mathbf{P}(y | x) \mathbb{P}[x_t = x, T_{x^*} > t \mid x_0 = x^*] \\ &= \mathbf{P}(y | x) \sum_{t=1}^{\infty} \mathbb{P}[x_t = x, T_{x^*} > t \mid x_0 = x^*] \\ &= \mathbf{P}(y | x) \mathbb{E} \left[\sum_{t=1}^{\infty} \mathbb{I}[x_t = x, T_{x^*} > t \mid x_0 = x^*] \right] \\ &= \mathbf{P}(y | x) \eta_{x|x^*} \end{aligned} \quad (3)$$

where $\eta_{x|x^*}$ denotes the expected number of visits to state x between two visits to state x^* . Furthermore:

- In (2), we used the fact that the event $\{T_{x^*} > t\}$ depends only on $\{x_i, 1 \leq i \leq t\}$, and thus it is independent of x_{t+1} ;
- In (3) we used the Markov property.

In order to perform the desired computation, we must try to find out the stationary distribution $\boldsymbol{\mu}^*$ of this chain:

$$\boldsymbol{\mu}^* = \boldsymbol{\mu}^* \mathbf{P} \Leftrightarrow \boldsymbol{\mu}^* = [0.0625 \quad 0.25 \quad 0.125 \quad 0.125 \quad 0.125 \quad 0.125 \quad 0.1875]$$

Now, on one hand, this chain is **irreducible** (since all states are accessible from all states - this stems from the fact the roads are bidirectional). On another hand, this chain is also **aperiodic**. This claim can be proven by successively computing powers of \mathbf{P} :

$$\begin{aligned}
\mathbf{P}^2 &= \begin{bmatrix} 0.25 & 0.0 & 0.25 & 0.25 & 0.25 & 0.0 & 0.0 \\ 0.0 & 0.625 & 0.0 & 0.0 & 0.0 & 0.125 & 0.25 \\ 0.125 & 0.0 & 0.2917 & 0.125 & 0.2917 & 0.1667 & 0.0 \\ 0.125 & 0.0 & 0.125 & 0.375 & 0.125 & 0.0 & 0.25 \\ 0.125 & 0.0 & 0.2917 & 0.125 & 0.2917 & 0.1667 & 0.0 \\ 0.0 & 0.25 & 0.1667 & 0.0 & 0.1667 & 0.4167 & 0.0 \\ 0.0 & 0.3333 & 0.0 & 0.1667 & 0.0 & 0.0 & 0.5 \end{bmatrix} \\
\mathbf{P}^3 &= \begin{bmatrix} 0.0 & 0.625 & 0.0 & 0.0 & 0.0 & 0.125 & 0.25 \\ 0.1562 & 0.0 & 0.2396 & 0.2188 & 0.2396 & 0.0833 & 0.0625 \\ 0.0 & 0.4792 & 0.0 & 0.0833 & 0.0 & 0.0625 & 0.375 \\ 0.0 & 0.4375 & 0.0833 & 0.0 & 0.0833 & 0.2708 & 0.125 \\ 0.0 & 0.4792 & 0.0 & 0.0833 & 0.0 & 0.0625 & 0.375 \\ 0.0625 & 0.1667 & 0.0625 & 0.2708 & 0.0625 & 0.0 & 0.375 \\ 0.0833 & 0.0833 & 0.25 & 0.0833 & 0.25 & 0.25 & 0.0 \end{bmatrix} \\
\mathbf{P}^4 &= \begin{bmatrix} 0.1562 & 0.0 & 0.2396 & 0.2188 & 0.2396 & 0.0833 & 0.0625 \\ 0.0 & 0.5052 & 0.0208 & 0.0417 & 0.0208 & 0.1302 & 0.2812 \\ 0.1198 & 0.0417 & 0.2448 & 0.151 & 0.2448 & 0.1667 & 0.0312 \\ 0.1094 & 0.0833 & 0.151 & 0.2448 & 0.151 & 0.0417 & 0.2188 \\ 0.1198 & 0.0417 & 0.2448 & 0.151 & 0.2448 & 0.1667 & 0.0312 \\ 0.0417 & 0.2604 & 0.1667 & 0.0417 & 0.1667 & 0.2604 & 0.0625 \\ 0.0208 & 0.375 & 0.0208 & 0.1458 & 0.0208 & 0.0417 & 0.375 \end{bmatrix} \\
\mathbf{P}^5 &= \begin{bmatrix} 0.0 & 0.5052 & 0.0208 & 0.0417 & 0.0208 & 0.1302 & 0.2812 \\ 0.1263 & 0.0417 & 0.2201 & 0.1914 & 0.2201 & 0.1146 & 0.0859 \\ 0.0104 & 0.4401 & 0.0208 & 0.0937 & 0.0208 & 0.0859 & 0.3281 \\ 0.0208 & 0.3828 & 0.0938 & 0.0417 & 0.0938 & 0.1953 & 0.1719 \\ 0.0104 & 0.4401 & 0.0208 & 0.0937 & 0.0208 & 0.0859 & 0.3281 \\ 0.0651 & 0.2292 & 0.0859 & 0.1953 & 0.0859 & 0.0417 & 0.2969 \\ 0.0937 & 0.1146 & 0.2187 & 0.1146 & 0.2187 & 0.1979 & 0.0417 \end{bmatrix}
\end{aligned}$$

For example, we can see that $\mathbf{P}^2(A | A) = 0.625$ and $\mathbf{P}^5(A | A) = 0.0417$ and thus they are both different from 0 (this can be confirmed from the fact that $A \rightarrow RP \rightarrow A$ and $A \rightarrow B \rightarrow F \rightarrow E \rightarrow C \rightarrow A$ are possible transitions in the chain). We are lead to conclude that $\gcd\{t \in \mathbb{N} : \mathbf{P}^t(A | A) > 0, t > 0\} = 1$ and, consequently, that the chain is aperiodic. Given the existence of a stationary distribution, this chain is also **positive**.

We are now equipped with the tools to perform our computation: given that the chain is **irreducible, aperiodic and positive**, the *Markov chain ergodic theorem* ensures us that the found stationary distribution is also **unique**. Since such stationary distribution μ^* can be written as ([1]):

$$\mu^*(x) = \frac{\eta_{x|x^*}}{\mathbb{E}[\mathbf{T}_{x^*} | \mathbf{x}_0 = x^*]}$$

Making the substitution $x = x^*$ in the expression above and taking into account that $\eta_{x^*|x^*} = 1$, we have that:

$$\mathbb{E}[\mathbf{T}_{x^*} | \mathbf{x}_0 = x^*] = \frac{1}{\mu^*(x^*)}$$

and so we have that

$$\eta_{x|x^*} = \mu^*(x) \mathbb{E}[\mathbf{T}_{x^*} | \mathbf{x}_0 = x^*] = \frac{\mu^*(x)}{\mu^*(x^*)}$$

Incorporating all this in equation (1), we have:

$$\mathbb{E} \left[\sum_{(x,y) \in E} T_{xy|RP} \right] = \sum_{(x,y) \in E} t_{xy} \mathbf{P}(y | x) \frac{\mu^*(x)}{\mu^*(x^*)}$$

Hence, if the truck leaves the recycling plant at 10am on Monday, it is expected to return to the plant after:

$$\begin{aligned} & t_{RPA} \mathbf{P}(A | RP) \frac{\mu^*(RP)}{\mu^*(RP)} + t_{ARP} \mathbf{P}(RP | A) \frac{\mu^*(A)}{\mu^*(RP)} + \\ & t_{AB} \mathbf{P}(B | A) \frac{\mu^*(A)}{\mu^*(RP)} + t_{BA} \mathbf{P}(A | B) \frac{\mu^*(B)}{\mu^*(RP)} + \\ & t_{AC} \mathbf{P}(C | A) \frac{\mu^*(A)}{\mu^*(RP)} + t_{CA} \mathbf{P}(A | C) \frac{\mu^*(C)}{\mu^*(RP)} + \\ & t_{AD} \mathbf{P}(D | A) \frac{\mu^*(A)}{\mu^*(RP)} + t_{DA} \mathbf{P}(A | D) \frac{\mu^*(D)}{\mu^*(RP)} + \\ & t_{CE} \mathbf{P}(E | C) \frac{\mu^*(C)}{\mu^*(RP)} + t_{EC} \mathbf{P}(C | E) \frac{\mu^*(E)}{\mu^*(RP)} + \\ & t_{BF} \mathbf{P}(F | B) \frac{\mu^*(B)}{\mu^*(RP)} + t_{FB} \mathbf{P}(B | F) \frac{\mu^*(F)}{\mu^*(RP)} + \\ & t_{DF} \mathbf{P}(F | D) \frac{\mu^*(D)}{\mu^*(RP)} + t_{FD} \mathbf{P}(D | F) \frac{\mu^*(F)}{\mu^*(RP)} + \\ & t_{EF} \mathbf{P}(F | E) \frac{\mu^*(E)}{\mu^*(RP)} + t_{FE} \mathbf{P}(E | F) \frac{\mu^*(F)}{\mu^*(RP)} \\ & = \frac{1}{0.0625} \left(30 \left(1 \times 0.0625 + \frac{1}{4} \times 0.25 \right) + 40 \left(\frac{1}{4} \times 0.25 + \frac{1}{2} \times 0.125 \right) + \right. \\ & \quad 55 \left(\frac{1}{4} \times 0.25 + \frac{1}{2} \times 0.125 \right) + 70 \left(\frac{1}{4} \times 0.25 + \frac{1}{2} \times 0.125 \right) + \\ & \quad 55 \left(\frac{1}{2} \times 0.125 + \frac{1}{2} \times 0.125 \right) + 80 \left(\frac{1}{2} \times 0.125 + \frac{1}{3} \times 0.1875 \right) + \\ & \quad \left. 70 \left(\frac{1}{2} \times 0.125 + \frac{1}{3} \times 0.1875 \right) + 20 \left(\frac{1}{2} \times 0.125 + \frac{1}{3} \times 0.1875 \right) \right) \\ & = 280 \text{ minutes} = \\ & = 4 \text{ hours } 40 \text{ minutes} \end{aligned}$$

Hence, it is expected for the truck to return to the plant at **2:40 pm on Monday**.

References

- [1] Francisco S. Melo. *Planning and Learning under Uncertainty*.