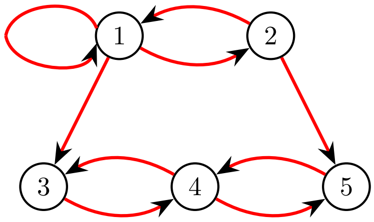
STK2130: Solution to the mandatory assignment 2024

Problem 1: a) The state diagram is (you may add transition probabilities if you want):



b) We see that states 1 and 2 communicate, but they do not communicate with the other states as it impossible to get from $\{3, 4, 5\}$ to $\{1, 2\}$. Hence $\{1, 2\}$ is a communication class, and it's transient as it's losing mass to $\{3, 4, 5\}$ that it cannot get back. The states 3, 4, 5 also communicate and form a communication class. Since a Markov chain with finite state space always has a recurrent communication class, $\{3, 4, 5\}$ must be recurrent.

c) Period is a class property, and hence it suffices to consider one state in each communication class. As it's possible to get from state 1 to itself in one step, state 1 (and hence state 2) has period 1. For the other class, we choose state 4 as our representative and observe that if we start at 4, we are in $\{3, 5\}$ at time 1, back at 4 at time 2, in $\{3, 5\}$ again at time 3 and so on. This means that we shall only be back at 4 at times divisible by 2, and hence 4 (and thus 3 and 5) has period 2.

d) Since states 1 and 2 are transient, we know that $\pi_1 = \pi_2 = 0$. To find π_3, π_4, and π_5, we consider the matrix equation $\pi = \pi P$:

$$(0, 0, \pi_3, \pi_4, \pi_5) = (0, 0, \pi_3, \pi_4, \pi_5) \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 \\ \frac{3}{4} & 0 & 0 & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} = (0, 0, \frac{1}{3}\pi_4, \pi_3 + \pi_5, \frac{2}{3}\pi_4)$$

This means that $\pi_3 = \frac{1}{3}\pi_4$, $\pi_4 = \pi_3 + \pi_5$, and $\pi_5 = \frac{2}{3}\pi_4$. These equations are solved by choosing π_4 freely, and then letting $\pi_3 = \frac{1}{3}\pi_4$ and $\pi_5 = \frac{2}{3}\pi_4$. In addition we need to have $1 = \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 = 0 + 0 + \frac{1}{3}\pi_4 + \pi_4 + \frac{2}{3}\pi_4 = 2\pi_4$, i.e. $\pi_4 = \frac{1}{2}$, and hence the stationary distribution is $(0, 0, \frac{1}{6}, \frac{1}{2}, \frac{1}{3})$.

e) If we start in state 4, the process will be back at 4 at all even times, i.e. $p_{44}^{(2n)} = 1$ for all n, and hence $p_{44}^{(n)} \nrightarrow \pi_4 = \frac{1}{2}$ as $n \to \infty$.

f) We offer two solutions to this problem:

Solution 1: If we use the method from the note "The Expected Number of Visits to Transient States", we first observe that since 1 and 2 are the only transient states, the transition matrix restricted to the transient states is

$$P_T = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \\ \frac{3}{4} & 0 \end{pmatrix}$$

We are interested in the matrix

$$S = (I - P_T)^{-1} = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} \\ \\ -\frac{3}{4} & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{12}{5} & \frac{4}{5} \\ \\ \frac{9}{5} & \frac{8}{5} \end{pmatrix}$$

According to the theory, s_{ij} is the expected number of visits to j when the process starts at i. Since we are starting at 1, the expected number of visits to 1 is $s_{11} = \frac{12}{5}$ and the expected number of visits to 2 is $s_{12} = \frac{4}{5}$.

Solution 2: It's not much harder to solve the problem from scratch: If we let v_1 and v_2 be the number of visits to 1 when we start from states 1 and 2, respectively, one step analysis tells us that

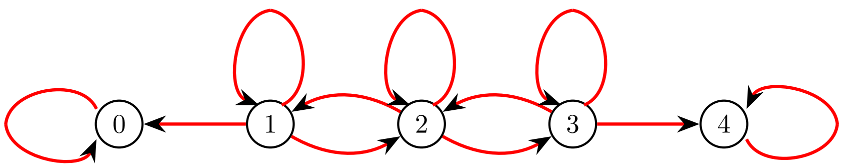
$$\begin{aligned} v_1 &= 1 + \frac{1}{3}v_1 + \frac{1}{3}v_2 \\ v_2 &= \frac{3}{4}v_1 \end{aligned}$$

We solve this system to get $v_1 = \frac{12}{5}$. If we similarly let u_1 and u_2 be the number of visits to 2 when we start from states 1 and 2, we get the equations

$$\begin{aligned} u_1 &= \frac{1}{3}u_1 + \frac{1}{3}u_2 \\ u_2 &= 1 + \frac{3}{4}u_1 \end{aligned}$$

Solving for u_1, we get $u_1 = \frac{4}{5}$.

Problem 2: a) The Markov chain will have 5 states corresponding to how many participants team A has on the top level, and we shall label the states 0,1,2,3,4 accordingly. The states 0 and 4 are absorbing as the game is over when we reach them. If i is one of the other states 1, 2, 3, it's possible to move to the neighboring states $i-1$ and $i+1$, and it's also possible to remain at i. Hence the state diagram is



To find the transition matrix, we observe that if we are in one of the middle states 2, 3, or 4, the conditions for moving to other states are:

1. Move from i to $i+1$: From the top level we must draw a player from team B, and from the bottom level we must draw one from team A.

2. Move from i to $i-1$: From the top level we must draw a player from team A, and from the bottom level we must draw one from team B.

3. Move from i to i: We must draw two players from the same team, either two from team A or two from team B.

Assume now that we are in state 2. To move to state 3, we need to pick a player from team B from the top level and one from team A from the bottom level. Both happens with probability $\frac{1}{2}$, and hence the combined probability is $\frac{1}{4}$. This means that $p_{23} = \frac{1}{4}$ and by symmetry $p_{21} = \frac{1}{4}$. This leaves $\frac{1}{2}$, and hence $p_{22} = \frac{1}{2}$ (you may want to double-check this by computing p_{22} directly).

Assume next that we are in state 3. In order to move to state 4, we need to pick a player from team B from the top level and one from team A from the bottom level. Both happens with probability $\frac{1}{4}$, and hence the combined probability is $\frac{1}{16}$. Thus $p_{34} = \frac{1}{16}$. If instead we look at the move from 3 to 2, we see that we need to pick a player from team A from the top level and one from team B from the bottom level. Both happens with probability $\frac{3}{4}$, and hence $p_{32} = \frac{3}{4} \cdot \frac{3}{4} = \frac{9}{16}$. This leaves a probability of $p_{33} = \frac{3}{8}$ to remain at 3.

By symmetry, we now see that $p_{10} = \frac{1}{16}$, $p_{11} = \frac{3}{8}$, $p_{12} = \frac{9}{16}$, and hence the transition matrix is

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{1}{16} & \frac{3}{8} & \frac{9}{16} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{9}{16} & \frac{3}{8} & \frac{1}{16} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

b) We want to find the number of steps needed to get to the absorbing states 0 and 4. If we delete the rows and columns belonging to these states, we get the reduced marix

$$R = \begin{pmatrix} \frac{3}{8} & \frac{9}{16} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{9}{16} & \frac{3}{8} \end{pmatrix}$$

We compute

$$(I-R)^{-1} = \begin{pmatrix} \frac{5}{8} & -\frac{9}{16} & 0 \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & -\frac{9}{16} & \frac{5}{8} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{44}{5} & 18 & \frac{36}{5} \\ 8 & 20 & 8 \\ \frac{36}{5} & 18 & \frac{44}{5} \end{pmatrix}$$

If $\mathbf{1}$ is the column vector with all entries equal to 1, the components of the vector

$$(I-R)^{-1}\mathbf{1} = \begin{pmatrix} 34 \\ 36 \\ 34 \end{pmatrix}$$

gives the expected number of stops needed to reach 0 or 4 starting at 1, 2, or 3, respectively. As we are starting at 2, the expected number of episodes is 36.

c) Let x, y, z be the probability that team A wins if the game is in state 1,2, and 3, respectively. One step analysis gives

$$\begin{aligned} x &= \frac{3}{8}x + \frac{9}{16}y \\ y &= \frac{1}{4}x + \frac{1}{2}y + \frac{1}{4}z \\ z &= \frac{9}{16}y + \frac{3}{8}z + \frac{1}{16} \end{aligned}$$

By symmetry $y = \frac{1}{2}$, and inserting this into the last equation, we get $z = \frac{9}{32} + \frac{3}{8}z + \frac{1}{16}$ which gives $z = \frac{11}{20}$, which is the winning probability when team A has three players on the top level.

d) Let $f\colon \{0,1,2,3,4\} \to \mathbb{R}$ be given by $f(0) = f(1) = f(2) = 0$, $f(3) = 100$, and $f(4) = 500$. Then the average number of points per game is

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(X_n),$$

and by the Ergodic Theorem,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(X_n) = \sum_{i=0}^{4} f(i)\pi_i,$$

where π is the stationary distribution.

To find the stationary distribution, note that the transition matrix for the new chain is

$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{16} & \frac{3}{8} & \frac{9}{16} & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 \\ 0 & 0 & \frac{9}{16} & \frac{3}{8} & \frac{1}{16} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

and hence the equation $\pi = \pi Q$ becomes

$$(\pi_0, \pi_1, \pi_2, \pi_3, \pi_4) = (\frac{1}{16}\pi_1, \pi_0 + \frac{3}{8}\pi_1 + \frac{1}{4}\pi_2, \frac{9}{16}\pi_1 + \frac{1}{2}\pi_2 + \frac{9}{16}\pi_3, \frac{1}{4}\pi_2 + \frac{3}{8}\pi_3 + \pi_4, \frac{1}{16}\pi_3)$$

This means that

$$\pi_0 = \frac{1}{16}\pi_1 \tag{1}$$

$$\pi_1 = \pi_0 + \frac{3}{8}\pi_1 + \frac{1}{4}\pi_2 \tag{2}$$

$$\pi_2 = \frac{9}{16}\pi_1 + \frac{1}{2}\pi_2 + \frac{9}{16}\pi_3 \tag{3}$$

$$\pi_3 = \frac{1}{4}\pi_2 + \frac{3}{8}\pi_3 + \pi_4 \tag{4}$$

$$\pi_4 = \frac{1}{16}\pi_3 \tag{5}$$

By symmetry we must have $\pi_4 = \pi_0$ and $\pi_3 = \pi_1$, and if we combine this with $\pi_0 = \frac{1}{16}\pi_1$, we get that $\pi_1 = \pi_3 = 16\pi_0$. Inserting this into equation (3), we get

$$\pi_2 = \frac{9}{16}\pi_1 + \frac{1}{2}\pi_2 + \frac{9}{16}\pi_3 = 9\pi_0 + \frac{1}{2}\pi_2 + 9\pi_0,$$

which gives $\pi_2 = 36\pi_0$. Hence $\pi = (\pi_0, 16\pi_0, 36\pi_0, 16\pi_0, \pi_0)$, and to get the components to sum to 1, we need to choose $\pi_0 = \frac{1}{70}$. This means that the stationary distribution is

$$\pi = \left(\frac{1}{70}, \frac{8}{35}, \frac{18}{35}, \frac{8}{35}, \frac{1}{70}\right)$$

(as we have taken som shortcuts, you may want to check that this distribution really is stationary).

Having found the stationary distribution, we can now return to the original problem:

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(X_n) = \sum_{i=0}^{4} f(i)\pi_i = 100 \cdot \frac{8}{35} + 500 \cdot \frac{1}{70} = \frac{1600+500}{70} = 30$$

Hence the average number of points won by the team per episode is 30.

Problem 3: a) If we count time in hours, the rates of the three waiting distributions are $\lambda_A = 4$, $\lambda_B = 3$, and $\lambda_C = 1$. The waiting time for the first bus to come is then exponentially distributed with rate $\lambda = \lambda_A + \lambda_B + \lambda_C = 4 + 3 + 1 = 8$. This means that the expected waiting time is $\frac{1}{8}$ hours, i.e. 7.5 minutes.

b) The probability that the first bus is from company C is $\frac{\lambda_C}{\lambda_A+\lambda_B+\lambda_C} = \frac{1}{8}$.

c) There are at least three different ways to solve this problem:

Solution 1: Let $N(t)$ be the number of buses that have arrived up to time t. Then N is a Poisson process with rate $\lambda = 8$, and if $N'(t)$ is the number of buses that have stopped by time t, N' is N "thinned" (in the sense of Section 2.4) by a factor $\frac{3}{4}$, and hence is a Poisson process with rate $\lambda' = \frac{3}{4}\lambda = \frac{3}{4} \cdot 8 = 6$. This means that the waiting time for the first bus to stop is exponentially distributed with rate 6 (and hence has mean $\frac{1}{6}$ hour $= 10$ minutes).

Solution 2: Let T be the waiting time for the first bus to arrive and let S be the waiting time for the first bus to stop. We know that T is exponentially distributed with rate $\lambda = 8$. To compute the density of S, note that if $r = \frac{3}{4}$ is the probability for each bus to stop, then the probability that the first bus to stop is bus number n equals $r(1-r)^{n-1}$. We know that the density of the arrival of bus number n is $\lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$, and hence the density of S is

$$f_S(t) = \sum_{n=1}^{\infty} r(1-r)^{n-1}\lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}$$

$$= r\lambda e^{-\lambda t}\sum_{n=1}^{\infty}\frac{((1-r)\lambda t)^{n-1}}{(n-1)!} = r\lambda e^{-\lambda t}e^{(1-r)\lambda t} = (r\lambda)e^{-(r\lambda)t}$$

which shows that S is exponentially distributed with rate $\lambda' = r\lambda$. In our case, $\lambda = 8$ and $r = \frac{3}{4}$, and hence $\lambda' = 6$. This means that you must expect to wait for 10 minutes for a bus that stops.

Solution 3: Since the bus arrivals are exponential (and hence memoryless), and the number of buses that have passed without stopping, doesn't influence whether future buses will stop or not, your expectations do not change with time, and hence the waiting time for buses that stop is also memoryless. As the exponential distributions are the only continuous, memoryless distributions, the waiting time for the buses that stop is also exponential. To find the rate, we compute the waiting time S for the first bus to stop (in minutes):

$$E[S] = \sum_{n=1}^{\infty} \left(\frac{1}{4}\right)^{n-1} \frac{3}{4} nE[T] = \frac{3}{4} E[T] \sum_{n=1}^{\infty} n \left(\frac{1}{4}\right)^{n-1} = \frac{3}{4} \cdot 7.5 \cdot \frac{1}{\left(\frac{3}{4}\right)^2} = 10$$

(here we have used the formula $\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}$). As the waiting time is 10 minutes $= \frac{1}{6}$ hour, the rate is 6.

d) Let $N'(t)$ be the number of buses that have stopped up till time t. Then $N'(t)$ is a Poisson process with rate $\lambda' = 6$, and $M = N'(\frac{1}{2})$ is Poisson distributed with rate 3. Hence $E[M] = 3$ and $\sigma_M = \sqrt{\mathrm{Var}(M)} = \sqrt{3}$. The probability that no more than one bus has stopped, is

$$P[M \leq 1] = P[M = 0] + P[M = 1] = \sum_{n=0}^{1} e^{-3} \frac{3^n}{n!} = e^{-3} + e^{-3} \frac{3}{1} = e^{-3} \cdot 4 \approx 0.1991,$$

i.e. about 20 %.