Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

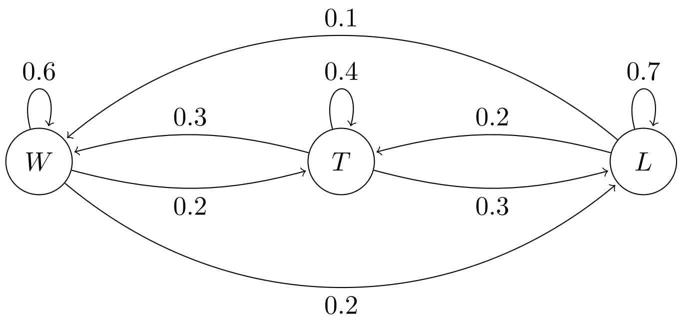
HS 2015 Prof. R. Wattenhofer / G. Bachmeier, S. Brandt, K.-T. Foerster, J. Seidel

Discrete Event Systems

Solution to Exercise Sheet 6

1 Soccer Betting

a) The following Markov chain models the different transition probabilities (W:Win, T:Tie, L:Loss):



b) The transition matrix P is

$$P = \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.2 & 0.7 \end{pmatrix} .$$

As you might have noticed, we gave redundant information here. You only need the information that the FCB lost its last game. Thus, the Markov chain is currently in the state L and hence, the initial vector is $q_0 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$. The probability distribution q_2 for the game against the FC Zurich is therefore given by

$$\begin{aligned} q_2 = q_0 \cdot P^2 = (q_0 \cdot P) \cdot P &= \begin{pmatrix} 0.1 & 0.2 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} 0.6 & 0.2 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.2 & 0.7 \end{pmatrix} \\ &= \begin{pmatrix} 0.19 & 0.24 & 0.57 \end{pmatrix} . \end{aligned}$$

(Note that q_0 must be a row vector, not a column vector.)

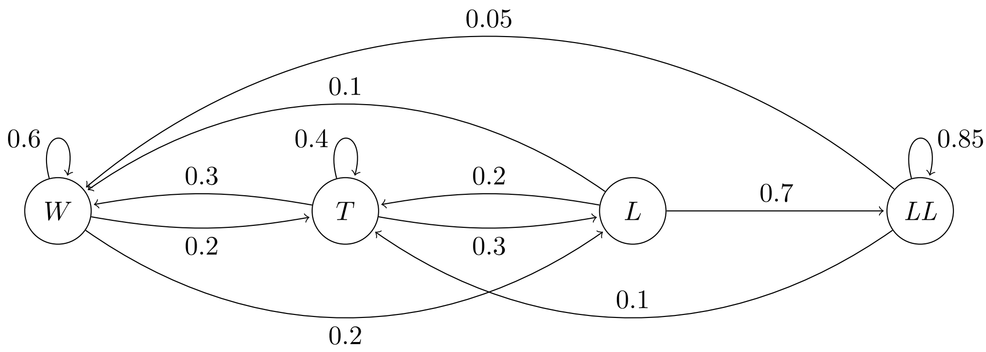
Hint: We exploited the associativity of the matrix multiplication to avoid having to calculate P^2 explicitly. This is usually a good "trick" to avoid extensive and error-prone calculations if no calculator is at hand (as for example in an exam situation ⌣̈).

Given the quotas of the exercise, the expected return for each of the three possibilities (W, T, L) calculates as follows.

$$\begin{aligned} \mathbf{E}[W] &= 0.19 \cdot 3.5 = 0.665 \\ \mathbf{E}[T] &= 0.24 \cdot 4 \quad = 0.96 \\ \mathbf{E}[L] &= 0.57 \cdot 1.5 = 0.855 \end{aligned}$$

Therefore, the best choice is not to bet at all since the expected return is smaller than 1 for every choice. If a “sales representative” of the Swiss gambling mafia were to force you to bet, you would be best off with betting on a tie, though.

c) The new Markov chain model looks like this. In addition to the three states W, T, and L, there is now a new state LL which is reached if the team has lost twice in a row.



The new transition matrix P is

$$P = \begin{pmatrix} 0.6 & 0.2 & 0.2 & 0 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0.1 & 0.2 & 0 & 0.7 \\ 0.05 & 0.1 & 0 & 0.85 \end{pmatrix} . \tag{1}$$

As the FCB has and lost its last two games, the Markov chain is currently in the state $q_0 = \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix}$. The probabilities for the game against the FC Zurich can again be computed as follows.

$$q_3 = q_0 \cdot P^2 = (q_0 \cdot P) \cdot P = \begin{pmatrix} 0.05 & 0.1 & 0 & 0.85 \end{pmatrix} \cdot \begin{pmatrix} 0.6 & 0.2 & 0.2 & 0 \\ 0.3 & 0.4 & 0.3 & 0 \\ 0.1 & 0.2 & 0 & 0.7 \\ 0.05 & 0.1 & 0 & 0.85 \end{pmatrix}$$

$$= \begin{pmatrix} 0.1025 & 0.135 & 0.04 & 0.7225 \end{pmatrix}$$

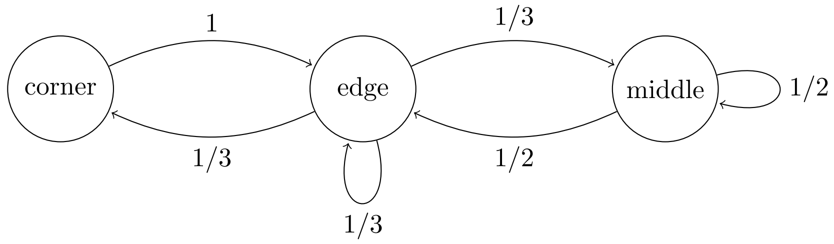
Finally, we can compute the expected profit for each of the three possible bets:

$$\begin{aligned} \mathbf{E}[W] &= 0.1025 \cdot 3.5 &&= 0.35875 \\ \mathbf{E}[T] &= 0.135 \cdot 4 &&= 0.54 \\ \mathbf{E}[L] &= (0.04 + 0.7225) \cdot 1.5 &&= 1.14375 \ . \end{aligned}$$

Now, the best choice is to bet on a loss. Clearly, the addition of the state LL worsens the situation for FCB.

2 Night Watch

a) Observe that the problem is symmetric, e.g., from all four corners, the situation looks the same, and the probability of being in a specific corner room is the same for all corners. The same holds for rooms at the border and for rooms in the middle. Thus, instead of using 16 states, we consider the following simplified Markov chain consisting of three states only:



The transition matrix M is given as follows.

$$M = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

To calculate the steady state probability, we have to calculate the eigenvector of M to the eigenvalue 1, that is solve the equation

$$v \cdot M = v$$

for v (Be careful to multiply M from the right side). Intuitively this means that if we have a state distribution v and applying the transition matrix does not change this distribution v, then v is the steady state distribution.

For $v = (c, e, m)$, we get the following system of linear equations from the above equation.

$$c = \frac{1}{3}e \qquad e = \frac{1}{3}e + \frac{1}{2}m + c \qquad 1 = c + e + m$$

Solving this equation system gives: $c = \frac{1}{6}$. The probability of being in a specific corner is therefore $\frac{1}{6} \cdot \frac{1}{4} = \frac{1}{24}$.

b) Since the two walks are independent, we have—according to the inclusion-exclusion principle (Einschluss-Ausschluss-Verfahren) –

$$\frac{1}{24} + \frac{1}{24} - \left(\frac{1}{24}\right)^2 \approx 0.082 \ .$$

3 PageRank

a) With $v = (1,1,1,1) \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}$ we get a PageRank vector v of $(1, 2, 2, 0)$. Notice that website v_4 is quite important for the PageRank, even though it is just a collection of links!

b) We first calculate $d_1 = 1, d_2 = 1, d_3 = 0, d_4 = 3$ and now get a PageRank vector of

$$v = (1,1,1,1) \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix} = (\frac{1}{3}, \frac{4}{3}, \frac{4}{3}, 0)$$

We now reduced the importance of the link—collection v_4, but still, we do not account for the fact that v_4 is not important at all—nobody recommends it!

c) The results of your iterations should look like this:

(i) $(\frac{1}{3}, \frac{4}{3}, \frac{4}{3}, 0)$

(ii) $(0, \frac{1}{3}, \frac{4}{3}, 0)$

(iii) $(0, 0, \frac{1}{3}, 0)$

(iv) $(0, 0, 0, 0)$

(v) $(0, 0, 0, 0)$

As you can see, when the Markov chain is not ergodic we run into problems. The $(1-\alpha)\cdot R$ component in the Google matrix M ensures that M is ergodic, and thus, due to the Ergodic Markov Chain theorem, the process would converge towards the unique stationary distribution. However, continuing these calculations including R might get a bit tedious if you are not a computer—so let us look at the other exercises now :-)

4 Probability of Arrival

The proof is similar to the one about the expected hitting time h_{ij} (see script). We express f_{ij} as a condition probability that depends on the result of the first step in the Markov chain. Recall that the random variable T_{ij} is the *hitting time*, that is, the number of steps from i to j. We get $Pr[T_{ij} < \infty \mid X_1 = k] = \Pr[T_{kj} < \infty] = f_{kj}$ for $k \neq j$ and $Pr[T_{ij} < \infty \mid X_1 = j] = 1$. We can therefore write f_{ij} as follows.

$$\begin{aligned} f_{ij} = \Pr[T_{ij} < \infty] &= \sum_{k\in S} \Pr[T_{ij} < \infty \mid X_1 = k] \cdot p_{ik} \\ &= p_{ij} \cdot \Pr[T_{ij} < \infty \mid X_1 = j] + \sum_{k\neq j} \Pr[T_{ij} < \infty \mid X_1 = k] \cdot p_{ik} \\ &= p_{ij} + \sum_{k\neq j} p_{ik} f_{kj} \end{aligned}$$

5 Basketball

a) This exercise is a good example to illustrate that most exercises allow several differing solutions.

Variant A. Let X be a random variable for the number of shots scored by Mario and X_i an indicator variable that the i-th shot scores. Then obviously $X = \sum_{i=1}^{n} X_i$ when n is the number of shots performed. The probability that the i-th attempt scores is p as given in the exercise. Hence, we can use linearity of expectation to obtain the expectation of X.

$$\mathbf{E}[X] = \mathbf{E}\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i=1}^{n} \mathbf{E}[X_i] = \sum_{i=1}^{n} p = n \cdot p$$

We want Mario to score m times.

$$\mathbf{E}[X] = n \cdot p = m \quad \Longleftrightarrow \quad n = \frac{m}{p}$$

Hence, Mario needs $\frac{m}{p}$ attempts to score m times. After these $\frac{m}{p}$ attempts, Mario has scored an expected m hits and he has missed expected $\frac{m}{p} - m$ times. Hence, he does an expected $10(\frac{m}{p} - m)$ push-ups in the game.

Variant B. We define a random variable X that counts the number of attempts until we miss for the first time. X is distributed as follows:

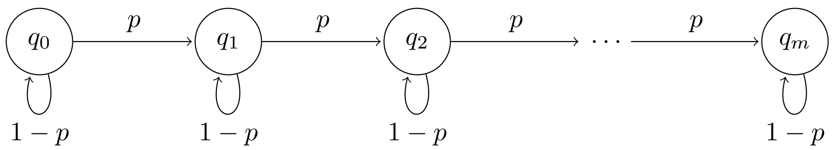
$$
\begin{aligned}
\Pr[X = 1] &= (1 - p) \\
\Pr[X = 2] &= p(1 - p) \\
&\ \vdots \\
\Pr[X = i] &= p^{i-1}(1 - p)
\end{aligned}
$$

We say that X is geometrically distributed with parameter $(1-p)$ or write $X \sim \mathrm{Geom}(1-p)$. The expected value of a geometrically distributed random variable with parameter α is $\frac{1}{\alpha}$.

$$\mathbf{E}[X] = \frac{1}{1-p} \quad .$$

Again, due to the linearity of the expected value we may think of the game as Mario scoring $\mathbf{E}[X] - 1$ hits, missing once, scoring the next $\mathbf{E}[X] - 1$ hits, missing again, and so forth until he scored a total of m hits. The question of how often Mario misses now translates to the question of how many series of $\mathbf{E}[X] - 1$ successful attempts he needs in order to score m times, and we get $10 \cdot \frac{m}{\mathbf{E}[X]-1} = 10 \cdot (\frac{m}{p} - m)$ push-ups in expectation.

Variant C (Markov Chain). The following Markov chain models Mario's game.



In a state q_i Mario has scored i hits. To learn the expected number of attempts until Mario has scored m hits we can simply compute the hitting time h_{0m} from q_0 to q_m.

$$
\begin{aligned}
h_{0m} &= 1 + \sum_{k \neq m} p_{0k} h_{km} = 1 + p_{00} h_{0m} + p_{01} h_{1m} \\
h_{0m} &= \frac{1 + p_{01} h_{1m}}{1 - p_{00}} = \frac{1 + p h_{1m}}{p} = \frac{1}{p} + h_{1m} \\
h_{1m} &= 1 + p_{11} h_{1m} + p_{12} h_{2m} \Longleftrightarrow h_{1m} = \frac{1}{p} + h_{2m} \\
h_{0m} &= \frac{1}{p} + h_{1m} = \frac{2}{p} + h_{2m} = \ldots = \frac{m}{p} + h_{mm} = \frac{m}{p}
\end{aligned}
$$

By subtracting the m successful attempts, we get an expected $\frac{m}{p} - m$ misses and hence Mario does $10(\frac{m}{p} - m)$ push-ups in expectation.

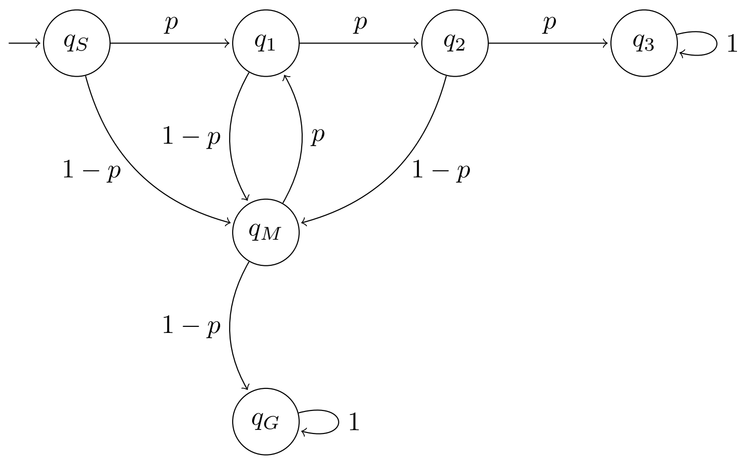
b) Each sequence of (at most m) throws where Luigi tries to score m times is called a *round*. A non-successful round is followed by push-ups.

Let X be a random variable for the number of rounds that Luigi has to perform until he hits m shots straight. The probability that Luigi scores m consecutive shots is p^m. Observe that X is geometrically distributed with parameter p^m (cf. Exercise 3a variant B) and hence

$$\mathbf{E}[X] = \frac{1}{p^m} \quad .$$

In the last round (which was successful), Luigi does not do any push-ups, hence we expect him to do $10 \cdot \left(\frac{1}{p^m} - 1\right)$ push-ups.

c) The following Markov chain models Trudy's game.



In state q_i Trudy has scored i hits in a row, in q_M she has missed once, in q_G she has missed twice in a row and gives up.

(i) We determine the probability f_{S3} of reaching the accepting state q_3 from the start state q_S.

$$\begin{aligned} f_{S3} &= p \cdot f_{13} + (1-p) \cdot f_{M3} \\ f_{13} &= p \cdot f_{23} + (1-p) \cdot f_{M3} \\ f_{23} &= p + (1-p) \cdot f_{M3} \\ f_{M3} &= p \cdot f_{13} \end{aligned}$$

$$\begin{aligned} f_{13} &= p^2 + (1-p)p^2 \cdot f_{13} + (1-p)p \cdot f_{13} \\ &= \frac{p^2}{1+p^3-p} = 0.4 \end{aligned}$$

$$\begin{aligned} f_{S3} &= p \cdot \frac{p^2}{1+p^3-p} + (1-p)p \cdot \frac{p^2}{1+p^3-p} \\ &= \frac{2p^3-p^4}{1+p^3-p} \\ &= 0.3 \end{aligned}$$

The probability that Trudy scores 3 times in a row is 0.3. The probability f_{SG} that she gives up is 0.7. This is because q_3 and q_G are the only absorbing states, i.e., all other states have probability mass of 0 in the steady state.

(ii) To get the number of push-ups we define a random variable Z that counts how often the system passes state q_M before either ending up in state q_3 or in state q_G. E.g., the probability $P[Z = 1]$ of passing q_M exactly once equals the probability of getting from q_S to q_M without being absorbed by q_3 and then ending up directly in q_G or q_3, i.e. $\Pr[Z = 1] = P_{SM} \cdot (P_{MG} + P_{M3})$ where P_{ij} is the probability of getting from q_i

to q_j without passing q_M on the way. Z has the following probability distribution:

$$\begin{aligned}
\Pr[Z = 0] &= 1 - P_{SM} \\
\Pr[Z = 1] &= P_{SM} \cdot (P_{MG} + P_{M3}) \\
\Pr[Z = 2] &= P_{SM} \cdot P_{MM} \cdot (P_{MG} + P_{M3}) \\
\Pr[Z = 3] &= P_{SM} \cdot P_{MM}^2 \cdot (P_{MG} + P_{M3}) \\
&\vdots \\
\Pr[Z = i] &= P_{SM} \cdot P_{MM}^{i-1} \cdot (P_{MG} + P_{M3})
\end{aligned}$$

The probability of passing q_M exactly i times equals the probability of getting from q_S to q_M and from q_M to q_M again $i-1$ times and then ending up directly in q_G or q_3. As the Markov chain is not too complicated we can compute the needed P_{ij} rather easily and get $P_{SM} = 1 - p^3$, $P_{MM} = p - p^3$, $P_{MG} = 1 - p$, and $P_{M3} = p^3$. The expected number of misses is

$$\begin{aligned}
\mathbf{E}[Z] &= \sum_{i=1}^{\infty} i \cdot \Pr[Z = i] \\
&= \sum_{i=1}^{\infty} i \cdot P_{SM} \cdot P_{MM}^{i-1} \cdot (P_{MG} + P_{M3}) \\
&= P_{SM} \cdot (P_{MG} + P_{M3}) \cdot \sum_{i=1}^{\infty} i \cdot P_{MM}^{i-1} \\
&= \frac{P_{SM} \cdot (P_{MG} + P_{M3})}{(1 - P_{MM})^2} \\
&= \frac{(1-p^3) \cdot (1 - p + p^3)}{(1 - p + p^3)^2} = \frac{1 - p^3}{1 - p + p^3} \\
&= \frac{1 - \frac{1}{8}}{1 - \frac{1}{2} + \frac{1}{8}} = \frac{7}{5} = 1.4.
\end{aligned}$$

Hence, Trudy does 14 push-ups in expectation.

Variant (for part (ii))**.** We already know that Trudy gives up with a probability 0.7. Let Z_i be the indicator random variable that is 1 if $X_i = q_M$ and 0 otherwise, so that $Z = \sum_{i=0}^{\infty} Z_i$. Note that f_{SG} can also be written as

$$0.7 = f_{SG} = \sum_{i=0}^{\infty} (\Pr[X_i = q_M] \cdot (1 - p)) ,$$

because q_G can only be reached from state q_M and each time Trudy is in q_M she gets to q_G with probability $1 - p$. Since Z_i is an indicator random variable, we know that $\Pr[X_i = q_m] = E[Z_i]$. Rearranging and using linearity of expectation, we obtain that

$$\frac{0.7}{1 - p} = \sum_{i=0}^{\infty} \mathbf{E}[Z_i] = \mathbf{E}[Z] .$$

Plugging in p yields that $\mathbf{E}[Z] = 1.4$, which means that the expected number of push-ups is 14.