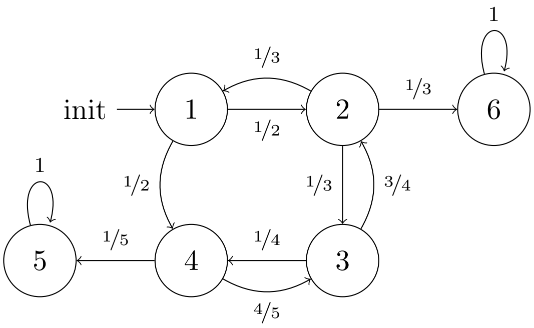
Exercise Sheet 2

General remarks:

- **Due date:** October 28th 12:30 (before the exercise class).
- Please submit your solutions via MOODLE. Remember to provide your matriculation number. It is necessary to hand in your solutions in groups of **three**. You may use the MOODLE forum to form groups.
- Solutions must be written in English.
- While we will publish sketches of exercise solutions, we do *not* guarantee that these sketches contain all details that are necessary to properly solve an exercise. Hence, it is recommended to attend the exercise classes.
- If you have any questions regarding the lecture or the exercise, please use the forum in MOODLE.
- Provide solutions to programming exercises in a separate plain text file. Please name the files according to the exercise they belong to (e.g. 'sheet01_ex1a.wppl'). All your programs are required to run without errors in the online interpreter available at `http://webppl.org`.

Exercise 1 (Reachability in Markov chains) $\boxed{\textbf{25P}}$

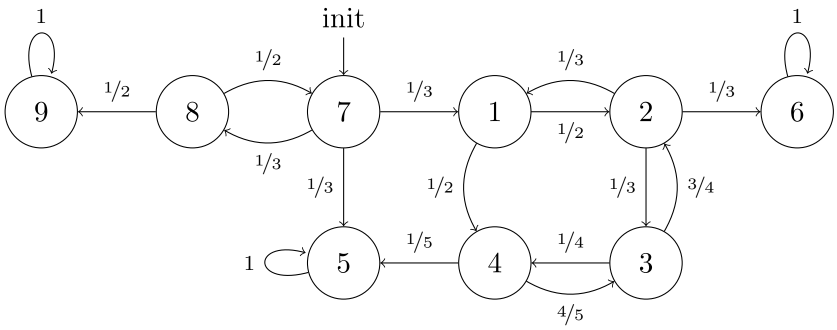
Consider the following Markov chain with state space $\Sigma = \{1, 2, 3, 4, 5, 6\}$ and goal states $G = \{6\}$:



(a) [5P] Compute the set $\Sigma_?$ as defined in the lecture.

(b) [5P] Set up the equation system $(\mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{b}$ defined in the lecture by giving the matrix $\mathbf{A}$ and the vector $\mathbf{b}$. You do not need to solve the system.

(c) [5P] We now extend the Markov chain from above as follows:



This time, let $G = \{5, 6\}$. Compute $Pr(\Diamond G)$ and justify your answer!

(d) [10P] A *sub-stochastic matrix* $\mathbf{M}$ is a square matrix with coefficients in $[0, 1]$ such that each row sum is *at most* 1. M is called *proper* if at least one row sum is *strictly smaller than 1*. Now let $D = (\Sigma, \sigma_I, \mathbf{P})$ be any Markov chain with goal set $\emptyset \neq G \subseteq \Sigma$ such that $\Sigma_? \neq \emptyset$. Prove or disprove: The matrix $\mathbf{A}$ is a proper sub-stochastic matrix.

Solution:

(a) Recall that $\Sigma_?$ is the set of states from which $G = \{6\}$ is reachable (excluding G itself):

$$\Sigma_? \;=\; Pre^*(\{6\}) \setminus \{6\} \;=\; \{1, 2, 3, 4\} \,.$$

In particular, 5 cannot reach $\{6\}$ because it is absorbing.

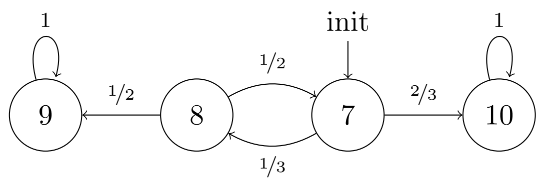
(b) The transition probability matrix $\mathbf{A}$ and the probabilities $\mathbf{b}$ to reach $\{6\}$ in one step are given by

$$\mathbf{A} \;=\; (\mathbf{P}(\sigma, \tau))_{\sigma,\tau \in \Sigma_?} \;=\; \begin{bmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/3 & 0 & 1/3 & 0 \\ 0 & 3/4 & 0 & 1/4 \\ 0 & 0 & 4/5 & 0 \end{bmatrix} \qquad \mathbf{b} \;=\; \begin{bmatrix} 0 \\ 1/3 \\ 0 \\ 0 \end{bmatrix}$$

The equation system $(\mathbf{I} - \mathbf{A}) \cdot \mathbf{x} = \mathbf{b}$ is thus

$$\begin{bmatrix} 1 & -1/2 & 0 & -1/2 \\ -1/3 & 1 & -1/3 & 0 \\ 0 & -3/4 & 1 & -1/4 \\ 0 & 0 & -4/5 & 1 \end{bmatrix} \cdot \mathbf{x} \;=\; \begin{bmatrix} 0 \\ 1/3 \\ 0 \\ 0 \end{bmatrix}$$

(c) We observe that the goal $G = \{5, 6\}$ is reached with probability 1 from all states $\{1, 2, 3, 4, 5, 6\}$ which can therefore be collapsed into a single goal state $G' = \{10\}$:



To compute the reachability probability in the simplified Markov chain, we use the equation system. It is $\Sigma_? = \{7, 8\}$, and

$$\mathbf{A} \;=\; \begin{bmatrix} 0 & 1/3 \\ 1/2 & 0 \end{bmatrix} \,, \qquad \mathbf{b} \;=\; \begin{bmatrix} 2/3 \\ 0 \end{bmatrix} \,.$$

Solving the system $(\mathbf{I} - \mathbf{A}) \cdot \mathbf{x} = \mathbf{b}$ yields $\mathbf{x} = \begin{bmatrix} 4/5 \\ 2/5 \end{bmatrix}$, thus $Pr(\Diamond G) = Pr(\Diamond G') = 4/5$.

(d) Let $\sigma \in \Sigma_? \neq \emptyset$. By definition of $\Sigma_?$, there is a path

$$\pi = \sigma, \ldots, \sigma', \sigma''$$

such that $\sigma'' \in G$. If we take such a π of minimum length, then $\sigma' \notin G$ and $\mathbf{P}(\sigma', \sigma'') > 0$, thus $\sigma' \in \Sigma_?$. Now consider the row of $\mathbf{A}$ that corresponds to σ', which is precisely the vector

$$\mathbf{P}(\sigma', \tau)_{\tau \in \Sigma_?} \ .$$

Since $\sum_{\tau \in \Sigma} \mathbf{P}(\sigma', \tau) = 1$ and $\mathbf{P}(\sigma', \sigma'') > 0$ and $\sigma'' \notin \Sigma_?$, it follows that

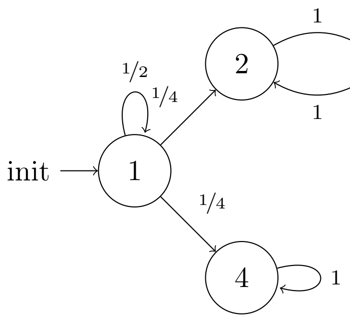
$$\sum_{\tau \in \Sigma_?} \mathbf{P}(\sigma', \tau) < 1 \ .$$

Therefore $\mathbf{A}$ is indeed a proper sub-stochastic matrix.

Remark: Proper sub-stochasticity of $\mathbf{A}$ implies that the matrix $\mathbf{I} - \mathbf{A}$ is non-singular in general.

Exercise 2 (Properties of Markov chains) **25P**

Consider the following Markov chain with state space $\Sigma = \{1, 2, 3, 4\}$:



(a) [4P] Classify each state of the Markov chains as either transient, positive recurrent or null recurrent!

Solution: Recall that in finite MC, a state is recurrent iff it is positive recurrent (i.e., null recurrence only plays a row in finite MC).

- State 1 is transient because the return probability is $f_1 = \frac{1}{2} < 1$.
- States $2, 3, 4$ are positive recurrent.

(b) [2P] Is the Markov chain irreducible? Justify your answer!

Solution:

- The MC is not irreducible, because e.g. state 2 cannot reach state 1.
- Alternatively, from part (a) we know that state 1 is transient, but Markov's theorem

assures that all states of irreducible MC are *recurrent*. Hence the Markov chain cannot be irreducible.

(c) [6P] Give all the irreducible components of the above Markov chain (Hint: There are 2). For each component say whether it is aperiodic and justify your claim.

Solution: The irreducible sub-MC are $\{2, 3\}$ and $\{4\}$.

- $\{2, 3\}$ is periodic with period $d = 2$. To see this note that for $\sigma \in \{2, 3\}$ we have $f_\sigma^{(n)} > 0$ iff n is even. Hence $d = 2$ is the largest integer satisfying

$$f_\sigma^{(n)} > 0 \implies (\exists k \geq 0 \colon n = k \cdot d) \ .$$

- The irreducible sub-MC $\{4\}$ satisfies $f_4^{(n)} = 1$ for all $n \geq 0$. Hence $d = 1$ which means that this components is aperiodic.

(d) [8P] Show that the Markov chain has infinitely many stationary distributions.

Solution: The transition probability matrix of the MC is

$$\mathbf{P} \ = \ \begin{bmatrix} 1/2 & 1/4 & 0 & 1/4 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \ .$$

Recall that the stationary distributions are exactly the non-negative (row-)vectors $\mathbf{x}$ that satisfy $\mathbf{x} = \mathbf{x} \cdot \mathbf{P}$ and $\sum_{\sigma \in \Sigma} \mathbf{x}_\sigma = 1$.
Now consider the family of vectors

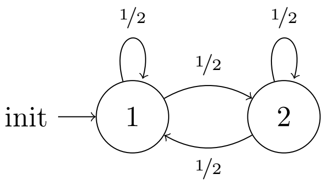
$$\mathbf{x}_p \ = \ \left(0, \ \frac{1}{2}p, \ \frac{1}{2}p, \ 1 - p\right)$$

for $p \in [0, 1]$. We have

$$\begin{aligned} \mathbf{x}_p \cdot \mathbf{P} \ &= \ \left(0, \ \frac{1}{2}p, \ \frac{1}{2}p, \ 1 - p\right) \cdot \begin{bmatrix} 1/2 & 1/4 & 0 & 1/4 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \ \left(0, \ \frac{1}{2}p, \ \frac{1}{2}p, \ 1 - p\right) \ . \end{aligned}$$

Furthermore, $0 + \frac{1}{2}p + \frac{1}{2}p + 1 - p = p + 1 - p = 1$.

(e) [5P] Give an example of an irreducible and aperiodic Markov chain with exactly 2 states and compute all its stationary distributions.



Solution: Since the Markov chain is irreducible and aperiodic, there is a *unique* stationary distribution which coincides with the limiting distribution. This distribution is $\rho = \left(\frac{1}{2}, \frac{1}{2}\right)$.

Exercise 3 (Rejection sampling) $\boxed{\textbf{25P}}$

(a) [10P] Write a WebPPL program that samples uniformly from a d-dimensional ball with radius 1 centered at the origin.

Solution:

```
var d = 10
var mcmc = false
var n = 1000

var model = function() {
     var x = repeat(d, function() {uniform({a:-1,b:1})});
     var squared = map(function(a){return a*a}, x)
     var squared_sum = sum(squared)
     condition(squared_sum < 1)
     return x
}

viz.auto(Infer({method: mcmc ? "MCMC" : "rejection", samples:n}, model))
```

(b) [5P] Compare (qualitatively) the time taken for inferring the posterior distribution using MCMC with Metropolis-Hastings kernel vs. rejection sampling both with parameters $d = 10$ dimensions and $n = 1000$ samples.

Solution: We observe that rejection sampling is slower than MCMC due to the "curse of dimensionality".

(c) [10P] Compute the probability that a sample gets accepted when using a uniform proposal distribution (on $[-1, 1]^d$) for $d \in \{10, 20, 40\}$ dimensions. Also give the expected total number of samples that have to be done for obtaining $n = 1000$ accepted samples. **Hints:**

- The acceptance rate of a sample is proportional to the ratio of volumes of the d-dimensional unit box and ball.
- The volume of a d-dimensional unit ball is $\frac{2\pi^{\frac{d}{2}}}{d \cdot \Gamma(\frac{d}{2})}$ where $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$.

Solution:

$$\begin{aligned}
d = 10: \quad & V_{\bigcirc}(d) = \frac{2\pi^5}{10 \cdot 4!} = \frac{\pi^5}{120} && \Pr(\text{accept}) = \frac{\frac{\pi^5}{120}}{2^{10}} \approx 0.00249 \\
d = 20: \quad & V_{\bigcirc}(d) = \frac{2\pi^{10}}{20 \cdot 9!} = \frac{\pi^{10}}{10 \cdot 9!} && \Pr(\text{accept}) = \frac{\frac{\pi^{10}}{20 \cdot 9!}}{2^{20}} \approx 2.46114 \cdot 10^{-8} \\
d = 40: \quad & V_{\bigcirc}(d) = \frac{2\pi^{20}}{40 \cdot 19!} = \frac{\pi^{20}}{20 \cdot 19!} && \Pr(\text{accept}) = \frac{\frac{\pi^{20}}{20 \cdot 19!}}{2^{40}} \approx 3.2785 \cdot 10^{-21}
\end{aligned}$$

For $n = 1000$ accepted samples, on average we need

$$\begin{aligned}
402,000 \quad & (d = 10), \\
4 \cdot 10^{10} \quad & (d = 20), \text{ and} \\
3 \cdot 10^{23} \quad & (d = 40) \text{ samples.}
\end{aligned}$$