GRAPHICAL MARKOV MODELS (WS2020)
2. SEMINAR

Assignment 1. (*Galton-Watson-Process – a population model*) Individuals of a certain population can have $n = 0, 1, 2\dots$ offspring at the end of their life. The corresponding probabilities are $c_0, c_1, c_2, \dots$. Let s_i denote the size of the population in the i-th generation.
a) Model the process as a Markov chain. Deduce a formula for the transition probabilities $p(s_i = k \mid s_{i-1} = m)$.
b****)** Calculate the extinction probability ρ_k, i.e. the probability that the population will eventually extinct if it starts with k individuals in the first generation.
Hints:

(1) Express ρ_k in terms of $\rho := \rho_1$.
(2) Try to find a functional relationship for ρ and the probabilities c_k, $k = 0, 1, 2, \dots$
(3) Analyse the resulting fix-point equation for ρ.

Let us consider the following standard Markov chain model for the next three assignments. The probability for sequences $s = (s_1, \dots, s_n)$ of length n with states $s_i \in K$ is given by:

$$p(s) = p(s_1) \prod_{i=2}^{n} p(s_i \mid s_{i-1}).$$

The conditional probabilities $p(s_i \mid s_{i-1})$ and the marginal probability $p(s_1)$ for the first element are assumed to be known.

Assignment 2.
a) Suppose that the marginal probabilities $p(s_i)$ for the states of the i-th element of the sequence are known for all $i = 2, \dots, n$. Then it is easy to compute all “inverse“ transition probabilities $p(s_{i-1} \mid s_i)$. How?
b) Describe an efficient algorithm for computing $p(s_i)$ for all $i = 2, \dots, n$.

Assignment 3. Suppose that there is a special state $k^* \in K$. We want to know how often this state appears on average in a sequence generated by the model. Describe an efficient method for computing this average.
Hint: Use the fact that the expected value of a sum of random variables is equal to the sum of their expected values.

Assignment 4. Let $A \subset K$ be a subset of states and let $\mathcal{A} = A^n$ denote the set of all sequences s with $s_i \in A$ for all $i = 1, \dots, n$. Find an efficient algorithm for computing the probability $p(\mathcal{A})$ of the event $\mathcal{A}$.

Assignment 5. According to Definition 1b of Sec. 1 of the lecture, any Markov chain model can be specified in the form

$$p(s) = \frac{1}{Z} \prod_{i=2}^{n} g_i(s_{i-1}, s_i)$$

with arbitrary functions $g_i\colon K^2 \to \mathbb{R}_+$ and the normalisation constant Z. Find an algorithm for computing the pairwise marginal probabilities $p(s_{i-1} = k, s_i = k')$ for all $k, k' \in K$ and all $i = 2, \ldots, n$ from the given functions g_i, $i = 2, \ldots, n$.

Assignment 6.* Suppose that a regular language $\mathcal{L}$ of strings over the finite alphabet Σ is described by a non-deterministic finite-state machine. Given a string $y \notin \mathcal{L}$, the task is to find the string $x \in \mathcal{L}$ with smallest Hamming distance to y, i.e.

$$x^* = \arg\min_{x \in \mathcal{L}} d_h(x, y),$$

where d_h denotes the Hamming distance. Construct an efficient algorithm for solving this task.