GRAPHICAL MARKOV MODELS
EXAM WS2015 (26P)

Assignment 1. (5p)
Let $S_0, S_2, \dots, S_{n-1}$ be a collection of K-valued random variables, where K is a finite set. Their joint probability distribution is a Markov model on a *cycle*

$$p(s) = \frac{1}{Z} \prod_{i=0}^{n-1} g_i(s_i, s_{i+1})$$

where indices $i+1$ are considered modulo n. The functions $g_i \colon K^2 \to \mathbb{R}_+$ are given and Z is a normalisation constant. Find an algorithm for searching the most probable realisation

$$s^* = \arg\max_{s \in K^n} p(s).$$

What complexity has it?

Assignment 2. (8p)
Let $S = (S_1, S_2, \dots, S_n)$ be a sequence of K-valued random variables, where K is a finite set. Their joint probability distribution is a Markov chain model given by

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

The transition probabilities $p(s_i \mid s_{i-1})$, $i = 2, 3, \dots, n$ and the probability distribution $p(s_1)$ of the first variable are known.

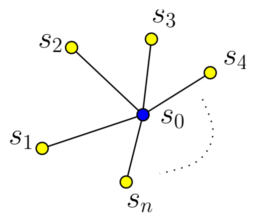
a) Given a subset $M \subset K$ of states and one of the variables S_i, the task is to compute the probability $p(s_i \in M)$, i.e. the probability that the variable S_i will take a value from M. Describe an efficient algorithm for computing this probability. What complexity has it?

b) Suppose now, it is known that the last variable S_n has taken the value $k^* \in K$. As before, the task is to compute the probability of S_i taking a value from M, but now conditioned on the additional information, i.e. $p(s_i \in M \mid s_n = k^*)$. Describe an algorithm for computing this probability. What complexity has it?

Assignment 3. (8p)
Let $S_0, S_2, \dots, S_n$ be a collection of K-valued random variables, where K is a finite set. Their joint probability distribution is a Markov model on a *star* given by

$$p(s) = p(s_0) \prod_{i=1}^{n} p(s_i \mid s_0) = \frac{1}{p(s_0)^{n-1}} \prod_{i=1}^{n} p(s_0, s_i)$$

(see figure below). The task is to learn the probabilities $p(s_0, s_i)$, $i = 1, \dots, n$ from i.i.d. training data $\mathcal{T} = \{\boldsymbol{s}^\ell \mid \ell = 1, 2, \dots, L\}$. Each example in $\mathcal{T}$ is a configuration $\boldsymbol{s} = (s_1, \dots, s_n)$ generated by the model in which, however, the value of s_0 is missing (not visible). Explain how to apply the Expectation Maximisation algorithm for this learning task.



Assignment 4. (5p)
Let (V, E) be an undirected graph and $\boldsymbol{s}$ denote a vertex labelling by labels from the finite set K, i.e $s_i \in K$ for all $i \in V$. The cost of a labelling $\boldsymbol{s}$ is

$$C(\boldsymbol{s}) = \sum_{i \in V} u_i(s_i) + \sum_{ij \in E} \mathbf{1}\{s_i \neq s_j\},$$

where the functions $u_i\colon K \to \mathbb{R}$ are given for all $i \in V$. The task is to find the labelling with minimal cost. Is it possible to solve the task approximately by applying the alpha-expansion algorithm? Justify your answer.