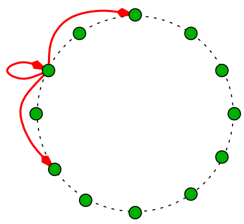
GRAPHICAL MARKOV MODELS
EXAM WS2013 (30+4P)

Assignment 1. (5p)
Consider a random walk on a discretised circle with positions denoted by $k \in \{0, 1, \ldots, K-1\}$ (see figure). The transition probabilities are defined as follows. The walker keeps staying at the current position k with probability α or it jumps to one of the two positions $(k \pm 2) \bmod K$ with probabilities $(1-\alpha)/2$. Find the conditions under which the corresponding Markov chain model is irreducible and a-periodic. Deduce its stationary distribution.



Assignment 2. (8p)
Suppose you have a Markov chain model for predicting the snow depth $s_t \in \{0, 1, \ldots, k_{\max}\}$ for a particular location based on daily measurements of temperature and precipitation[1] x_t, where the transition probabilities $p(s_t \mid s_{t-1}; x_t)$ are parametrised by the measurements. Your customer is however not interested in snow depth. All she wants to know, are binary valued snow cover values $c_t \in \{0, 1\}$. You have agreed to use the Hamming distance $d(c, c') = \sum_{t \in T} |c_t - c'_t|$ for loss.
a) Deduce the optimal Bayes decision for the sequence $c_1, \ldots, c_T$ given the measurements and the initial snow depth s_1.
b) Explain an efficient algorithm for solving this task.

Assignment 3. (8p)
Suppose you want to learn the transition probabilities of a homogeneous Markov chain model

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

with integer valued states $s_i \in \mathbb{Z}$. It is assumed that these transition probabilities depend on the differences $d_i = s_i - s_{i-1}$ only and are zero if $|d_i| > d_{\max}$.
a) Explain how to learn (estimate) the transition probabilities from an given i.i.d. sample $\mathcal{T}$ of sequences s.
b) Let us suppose, that the data sequences x given for learning are transformed by $x_i = \lfloor s_i/2 \rfloor$ due to some "resolution deficiency". The brackets $\lfloor\rfloor$ denote the integral part (floor). How would you learn the transition probabilities in such a situation?

[1]precipitation = srážky

Assignment 4. (8p)
Consider a (Min,+)-problem for K valued labellings $s\colon V \to K$ of a graph (V, E)

$$F(s) = \sum_{i \in V} u_i(s_i) + \sum_{ij \in E} u_{ij}(s_i, s_i) \to \min_{s \in K^V} .$$

Suppose we want to extend the objective function to $F'(s) = F(s) - \alpha N(s)$, where $N(s)$ "counts" the number of labels appearing in the labelling s, i.e.

$$N(s) = |\{k \in K \mid \exists i \in V \colon s_i = k\}|.$$

The extended objective function is not any more a sum of functions of arity two or less.
a) Find an equivalent representation of the extended optimisation task, such that the objective function is a sum of arity two functions (at most) by introducing auxiliary nodes (variables) and edges. (Hint: the label set for the auxiliary variables can be different from K).
b) Suppose that the initial (Min,+)-problem was submodular w.r.t. some ordering of K. Is the extended (Min,+)-problem constructed by you still submodular?

Assignment 5. (5p)
Consider a Markov chain model for K-valued sequences s

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

Suppose that $\sim$ is an equivalence relation defined on the label set K and $x = \pi(s)$ denotes the projection of the sequence s to a sequence x of equivalence classes $K/{\sim}$. One might be tempted to think that the related probability distribution

$$p(x) = \sum_{s \colon \pi(s) = x} p(s)$$

is also a Markov chain model. Is that true? (Hint: considering chains with length three should be sufficient.)