STATISTICAL MACHINE LEARNING (WS2016)
SEMINAR 5

Assignment 1. See Assignment 5 from the previous seminar.

Assignment 2. Consider a Hopfield network with n binary valued units $y_i = \pm 1$, $i = 1, \ldots, n$. Given $m < n$ binary patterns $\boldsymbol{y}^\ell \in \{-1, +1\}^n$, we want to store them as local minima of the network energy. Show that the Hebb-rule, which defines the weights of the network by

$$w_{ij} = \frac{1}{m} \sum_{\ell=1}^{m} y_i^\ell y_j^\ell \text{ for all } i \neq j,$$

is sufficient to store them as local minima if the patterns are nearly orthogonal, i.e.

$$|\langle \boldsymbol{y}^\ell, \boldsymbol{y}^k \rangle| = |\sum_{i=1}^{n} y_i^\ell y_i^k| \leqslant 1 \text{ for all } \ell \neq k$$

Hint: Use the fix-point condition for local minima $\boldsymbol{y}$ of the energy

$$y_i = \operatorname{sign}\bigl(\sum_j w_{ij} y_j\bigr)\ \forall i = 1, \ldots, n.$$

Assignment 3. A standard sudoku on a 9x9 field has approximately 6.67×10^{21} solutions. Consider it as a graph labelling problem for a graph (V, E) with 81 nodes. Each node can be labelled by one of the nine labels from $K = \{1, 2 \ldots, 9\}$. Let us denote a labelling by $\boldsymbol{y} = (y_1, \ldots, y_{81})$.
a) Define a set of edges E and find a function $g\colon K^2 \to \mathbb{R}$ such that the energy function

$$F(\boldsymbol{y}) = \sum_{\{i,j\} \in E} g(y_i, y_j)$$

attains global minima at labellings $\boldsymbol{y}$ representing valid sudoku solutions.
b) Recall that a standard sudoku problem requires to find the labels y_i, $i \in V \setminus M$, given the labels y_i for some subset of nodes $M \subset V$. We will consider a generalised sudoku problem of the following form. Suppose you are given confidence values $0 \leqslant \alpha_i(k) \leqslant 1$ for labels $k \in K$ in each node $i \in V$. The task is to find a labelling (i.e. solution) with highest total confidence, i.e

$$A(\boldsymbol{y}) = \sum_{i \in V} \alpha_i(y_i) \to \max_{\boldsymbol{y}}$$

subject to the constraint that $\boldsymbol{y}$ must be a valid sudoku labelling. Generalise the function $F(\boldsymbol{y})$ such that its global minima represent the optimal solutions of this sudoku problem.