Neuronale Netze (SS 2002), 22.4.

Perceptron for non linearly separable data (continued):

- **Hitting set problem:** Given a set $S = \{s_1, \ldots, s_n\}$, a set of subsets $C = \{c_1, \ldots, c_m\}$ with $c_i \subset S$, $k \in \mathbb{N}$, does there exist k points in S such that every c_i is hit by at least one of the k points

  (In real life: Find a fixed number of representatives such that each group of interest is covered.)

  We have shown that the problem is NP-complete via reduction from SAT. This holds even under the restriction $|c_1| = |c_2| = \ldots$.

- **Perceptron training in the presence of errors:** Given an input dimension n, a training set $P \subset \{0,1\}^n \times \{0,1\}$, and $k \in \mathbb{N}$. Decide whether a perceptron can be found which has at most k errors on P.

  We have shown that this problem is NP-complete via a (complicated) reduction from the hitting set problem.

  $\Rightarrow$ optimum training of a perceptron is not that easy in principle!

Rosenblatt-Perceptron – an easy extension of the perceptron such that we can still use the perceptron learning algorithm for training:

- The Rosenblatt-perceptron computes a function

  $$f : \{0,1\}^{n \times m} \to \{0,1\}, \vec{x} \mapsto p(f_1(\vec{x}), \ldots, f_M(\vec{x}))$$

  where p is a standard perceptron and f_i are **masks/receptors** implementing some function.

  Idea: The f_i implement local problem dependent feature extraction, e.g. smoothing, line extraction, $\ldots$. The f_i are fixed, only p is trained accordingly. The hope is, that for most realistic functions f simple standard receptors f_i can be found such that the problem becomes linearly separable.

- If there is no restriction on the receptors f_i, every function f can be implemented with a Rosenblatt-perceptron – just choose $f_1 = f$.

- A receptor f_i has a limited diameter k, iff indices (j,l) exist such that f_i depends only on the input coefficients

$$\begin{array}{lll} x_{(j,l)} & \cdots & x_{(j,l+k-1)} \\ \vdots & & \vdots \\ x_{(j+k-1,l)} & \cdots & x_{(j+k-1,l+k-1)} \end{array}$$

- Hope: standard functions f can be made linearly separable only applying diameter-restricted receptors.

  Note: applying some filters for preprocessing in vision (smoothing, line-extraction, ...) corresponds to diameter-restricted masks. This is part of the standard preprocessing for real problems in vision – although the problems become unfortunately not linearly separable.

- A Rosenblatt-perceptron with diameter-restricted receptors cannot recognize connected patterns!

  $\Rightarrow$ Nice try, but unfortunately not enough!