Neural Networks

8 June 1998

***Grading*:** In total you can get 100 points. You get 10 points free. Grade = number of points/10.

Problem A (30 points) *Perceptron*

1 (5 p) Describe the perceptron training algorithm. Formulate the update rule for a single-layer perceptron.

2 (5 p) Describe the pocket algorithm (any variant of it).

3 (5 p) Define the concept of linear separability.

4 (5 p) Compare both algorithms (perceptron and pocket) with respect to the following criteria: convergence, speed, robustness (by 'robustness' we mean here the following property: the longer the network is trained the better accuracy is achieved).

5 (3 p) Let us consider two sets of points in R^2:

$A=\{(x,y) \mid x^2+y^2<144, x, y \text{ are integer}\}$

and

$B=\{(x,y) \mid x^2+y^2>144, x, y \text{ are integer}\}$.

Are these two sets lineary separable? Justify your answer.

6 (7 p) Let us consider similar sets of sets of points in R^4:

$A=\{(x,y,z,t) \mid x^2+y^2<144, z=x^2, t=y^2, x, y, z, t \text{ are integer}\}$

and

$B=\{(x,y,z,t) \mid x^2+y^2>144, z=x^2, t=y^2, x, y, z, t \text{ are integer}\}$.

Are these two sets lineary separable? Justify your answer.

Problem B (25 points) *Backpropagation*

Let us consider a feed-forward network shown in figure 1. Here we assume that the hidden units have no bias and that the activation function is given by $S(net)=net^3$.
(Let us recall that $(x^3)'=3x^2$).

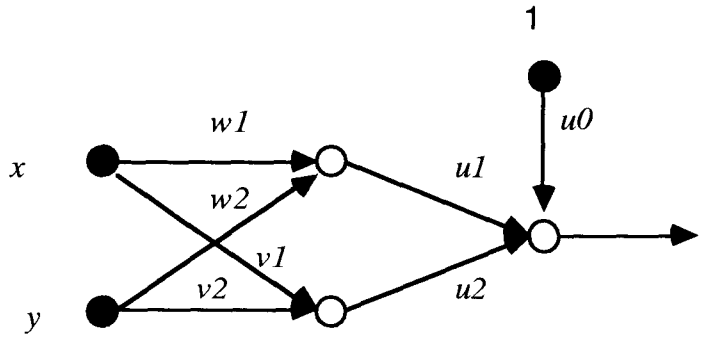

Figure 1. A 2:2:1 feed-forward network with 7 weights.

a) (5 p) Express the output of this network as a function of nine variables: x, y, w_1, w_2, v_1, v_2, u_0, u_1, u_2 (write a formula).

b) (10 p) Formulate update rules for weights u_0, u_1 and w_1. You don't have to derive these rules–just apply the generalized delta rule used in the backpropagation algorithm.

c) (5 p) Can this network be trained to solve the XOR problem? Briefly justify your answer.

d) (5 p) Let us suppose that the training parameter is set to a negative value (e.g. -0.5). How would it affect the training process? Would it converge? Would the result be a local minimum?

Problem C (15 points) *Constructive Algorithms*

a) (10 p.) Describe in detail one (and only one) of the following 2 algorithms:

- the upstart algorithm
- the cascade correlation algorithm

b) (5 p.) Which of these two algorithms is not suitable for function approximation? Justify your answer.

Problem D (20 points) *Overview*

Let us consider the following 5 types of networks:

a) winner-take-all networks (WTA networks)

b) feed-forward networks (BP networks)

c) Radial Basis Function networks (RBF networks)

d) Self-organizing maps (Kohonen networks)

e) Hopfield networks

Characterize each of these networks with respect to the following aspects:

1) types of problems that can be solved with these networks: classification, function approximation, data clustering,

2) type of learning algorithm: supervised/unsupervised/mixed,

3) speed of training,

4) the amount of training parameters (list them).

For each network give an example of its application.