UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

B.Sc. Examination 2004

COMPUTING AND INFORMATION SYSTEMS

IS53002A (CIS311) Neural Networks

Duration: 2 hours 15 minutes

Date and time:

- *Full marks will be awarded for complete answers to THREE questions. Do not attempt more than THREE questions on this paper. Although each question carries 25 marks, and therefore the result from three questions sums up to 75 marks, the final result will be additionally scaled to 100.*

- *Electronic calculators may be used. The make and model should be specified on the script. The calculator must not be programmed prior to the examination. Calculators which display graphics, text or algebraic equations are not allowed.*

THIS EXAMINATION PAPER MUST NOT BE

REMOVED FROM THE EXAMINATION ROOM

Question 1.

a) Explain briefly the following conceptual differences between the biological neural system and the conventional (von Neumann) computer system: processing device, memory, computing capacity and reliability. **[8]**

b) Which three main computational abilities of the biological neural networks are simulated by artificial neural networks? **[6]**

c) Explain how classification to K (K>2) classes can be performed using a single layer Perceptron network which is actually a linear discriminating function? **[7]**

d) Consider a neural network with a hidden layer having 2 nodes and an output layer having 1 node. Write down the expression for calculating the total number of weights (including biases) in this neural network. **[4]**

Question 2.

a) Determine the output of a single-layer Perceptron with two inputs, having weights w_1=0.25
and w_2=-0.5., when given an input vector (1,1) for each of the following activation functions:

i) linear activation function without bias (i.e. without threshold); **[2]**

ii) sigmoidal activation function; **[2]**

iii) thresholded activation function using bias (i.e. using threshold) 0.33. **[2]**

b) A single-layer Perceptron network with two inputs and a bias input is given. Train this single-layer network with the gradient-decent rule using the following sequence of examples (where x_0, x_1, x_2 are inputs and y is output):

	x_0	x_1	x_2	y
1)	1	-1	-1	1
2)	1	-1	1	1
3)	1	1	-1	-1
4)	1	1	1	1

Consider the following initial weights: (w_0, w_1, w_2) = (-0.2, 0.1, 0.3),
and learning rate η=0.5.

i) Demonstrate how the weights are updated after each example is presented (in incremental fashion); **[12]**

ii) Demonstrate how the weights are updated after all examples are presented (in batch fashion). **[7]**

Question 3.

a) Which are four main advantages of probabilistic neural networks in comparison with multilayer neural networks. **[4]**

b) Describe two characteristics of probabilistic neural networks that make them as useful as multilayer neural networks. Explain them briefly. **[6]**

c) Give the formula for processing the outputs from pattern nodes, associated with a given class,
used in the summation nodes of probabilistic neural networks. **[4]**

d) What will be the classification on the testing example: `(0.31 0.42 0.53 0.64)` calculated by the probabilistic network using the following four training vectors:

```
[0.5 0.45 0.4 0.5] positive
[0.5 0.44 0.3 0.6] positive
[0.2 0.41 0.3 0.7] negative
[0.1 0.43 0.4 0.8] negative
```

assuming that the squared spread is $\sigma^2=1$. **[11]**

Question 4.

a) Determine the number of inputs and the number of hidden nodes of a radial basis function
neural network whose weights are trained using the equation: $\mathbf{w} = (\mathbf{\Phi}^T \mathbf{\Phi} + \mathbf{I})^{-1} \mathbf{\Phi}^T \mathbf{y}$
where: **I** is the identity matrix, **y** is the output given with the examples $\mathbf{x}_e = (x_1, x_2, x_3)$,
and **Φ** is the design matrix of radial basis functions defined as follows:

$$\mathbf{\Phi} = \begin{matrix} \varphi_{11} & \varphi_{12} & \varphi_{13} & \varphi_{14} \\ \varphi_{21} & \varphi_{22} & \varphi_{13} & \varphi_{24} \\ \dots & \dots & \dots & \dots \\ \varphi_{N1} & \varphi_{N2} & \varphi_{13} & \varphi_{N4} \end{matrix}$$

i) Draw the architecture of this radial basis function neural network. **[5]**

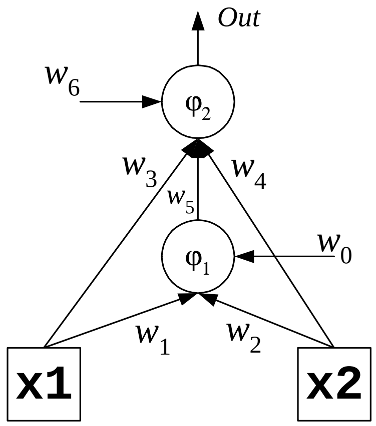
ii) Explain what is the purpose of each layer of this radial basis function network and what is computed at each layer? **[8]**

b) Design a simple radial basis function neural network for learning the examples of the boolean XOR function. Suggest which centers and spreads can be considered for accurate learning of the XOR examples. **[9]**

c) Which are the three most widely used approaches to finding the centres and variances for the radial basis function network training algorithm? **[3]**

Question 5.

A multilayer neural network with two nodes: one hidden and one output using sigmoidal activations
is given. There are two inputs to the network: (x_1, x_2), and seven weights as illustrated in the picture
below. Both of the hidden neurons have bias connections which are clamped at 1.



Perform training of this neural network using the backpropagation algorithm using zero momentum
and learning rate parameter 1. Consider the following initial weights:

w_0 = -1.1	w_1 = 0.5	w_2 = 0.5	
w_3 = 0.6	w_4 = 0.7	w_5 = -1.8	w_6 = -0.2

Conduct training with the following training vector:

x_1	x_2	y
0	1	1

Show the modification of each weight and the bias after one iteration of the backpropagation
algorithm performed with this training vector. Give the output, the error derivatives β (beta),
the weight updates *delta-w*, and finally the modified weights. **[25]**