Answer THREE questions.

1.

a) Discuss the ways in which current artificial neural networks differ from their biological counterparts. Consider both the operation of individual neurons and the organisation of neurons within the network. In what ways might future artificial neural networks benefit from incorporating further biological principles?

[15 marks]

b) A single McCullough-Pitts neuron may be used to perform many Boolean functions. By choosing appropriate weights and thresholds, demonstrate the neural implementation of

(i) AND
(ii) OR
(iii) NOT

[12 marks]

c) Explain with the aid of a suitable diagram why a single McCullough-Pitts neuron cannot implement the XOR function.

[6 marks]

TURN OVER

2.

a) Explain why training algorithms based on a process of gradient descent, such as error backpropagation, can be vulnerable to trapping in **local minima**. Discuss ways in which the training process may be modified in order to make this less likely to occur.

[10 marks]

b) Explain what is meant by **overtraining** and how this problem may be handled by the use of a **training test set** ('validation set').

[7 marks]

c) In a real-world application such as financial time series prediction, which is likely to cause the greater difficulties, overtraining or trapping in local minima? Why?

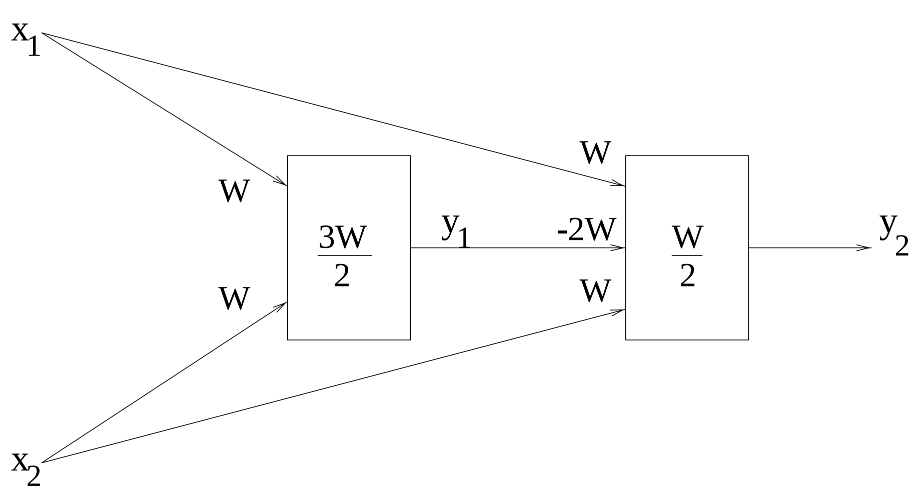
[4 marks]

d)

(i) Explain the role of a **feature detector** in a trained multilayer perceptron network.

[4 marks]

(ii) Consider the multilayer perceptron network below, where each of the processors has a real-valued output $y_i \in [0, 1]$, computed using a sigmoidal firing function. The quantities inside the boxes are the thresholds of the units, and those attached to the input lines are the weights associated with those inputs (where W is a large positive constant).



What function of its binary inputs does the network compute? What feature of the input does the hidden unit detect?

[8 marks]

3.

a) What quantities play the roles of ’input’ and ’output’ in a recurrent network of the Hopfield type?

[4 marks]

b) In what sense does the Hopfield net utilise a process of gradient descent?

[4 marks]

c) Explain why a multiplicity of local minima are desirable for a Hopfield network, whilst being undesirable for a multilayer perceptron network trained with error back-propagation.

[6 marks]

d) It is desired to store a set of P N-bit binary patterns

$$\underline{x}^{(p)} = (x_1^{(p)}, x_2^{(p)}, ..., x_N^{(p)}) \quad p = 1..P$$

in an N-node Hopfield net.

Write down the *simplest* storage prescription for the $\frac{N}{2}(N-1)$ weights w_{ij} and N thresholds s_i which will achieve this.

[3 marks]

e) It is desired to store the 3-bit binary patterns (0,0,1) and (1,0,0) in a 3-node Hopfield net.

Use the simple storage prescription of d) to obtain weights and thresholds which will achieve this. Draw a state transition diagram for the system, labelling the energy levels of the system, and each transition with its probability.

[10 marks]

f) Comment on the CAM properties displayed by your solution in e). Suggest a modified storage rule that might improve the quality of the solution, and explain why.

[6 marks]

4.

a) "Any problem that can be solved by a supervised learning rule can also be solved by reinforcement training". Justify this statement.

[7 marks]

b) Given the truth of the above assertion, why is reinforcement training not generally used as a substitute for supervised learning? Under what special circumstances might one choose to use reinforcement training to solve a supervised learning problem?

[7 marks]

c) The A_{RP} neuron is a 2-action stochastic learning automaton. The activation a_i of neuron i is a function of the context (input) vector $\underline{x}$:

$$a_i = \sum_{j=0}^{N} w_{ij} x_j \qquad \text{(bias input } x_0 = 1\text{)}$$

(i) Write down a stochastic **firing rule** for the A_{RP} neuron, as a function of its activation a_i

[4 marks]

(ii) Write down an **update rule** for the parameters w_{ij} in the presence of an externally generated reinforcement signal r. In what way does the ’penalty parameter’ λ in this expression affect the behaviour of the neural automaton?

[7 marks]

(iii) Show by considering binary reinforcement signals $r \in \{0, 1\}$ as well as binary actions that for any action/reinforcement combination the above rule implements an effective strategy for dealing with an initially unknown and unpredictable environment.

[8 marks]

CONTINUED

5.

a) Give a careful definition of each of the following terms:

(i) self-organisation
(ii) topographic map
(iii) dimensional reduction
(iv) winning node (in a Kohonen map)

[12 marks]

b) Write down the Kohonen update rule for the change in weights to a winning node and its neighbours in an m×m output grid. Define the **neighbourhood size** d(t) on such a grid. Explain how (and why) the neighbourhood size and training rate must be decremented in the course of training.

[6 marks]

c) Assuuming that the weight vectors $\underline{w}_i$ and input vectors $\underline{x}$ in a Kohonen net are *normalised*

$$\sum_j w_{ij}^2 = 1, \qquad \sum_j x_j^2 = 1$$

explain how a process of **local excitation** and **lateral inhibition** could be used to pick out a winning node and its neighbours. How could neighbourhood size be modified in this more biologically plausible version of the Kohonen algorithm?

[7 marks]

d) A neural network is proposed which utilises a 2-stage training process. In the first stage the raw data is fed into a single-layer Kohonen net which is then trained to completion. In the second stage the outputs of the Kohonen net are fed into a multilayer perceptron, which is trained using error backpropagation.

What do you believe is the rationale behind the construction of a hybrid network of this type? To what kinds of problems might such a network be applied? Do you think this approach would be successful?

[8 marks]

END OF PAPER