UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 1996

BEng Honours Degree in Computing Part III
MEng Honours Degrees in Computing Part IV
BSc Honours Degree in Mathematics and Computer Science Part III
MSc Degree in Foundations of Advanced Information Technology
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the
Diploma of Membership of Imperial College
Associateship of the City and Guilds of London Institute
Associateship of the Royal College of Science

PAPER 3.89 / 4.89

NEURAL NETS

Friday, May 3rd 1996, 10.00 - 12.00

Answer THREE questions

For admin. only: paper contains
4 questions
3 pages (excluding cover page)

1. a Give a high level pseudo-code description of a typical genetic algorithm and discuss those features which most influence various aspects of the algorithm's behaviour.

b i) In applying a typical genetic algorithm to a particular problem identify two design considerations likely to be critical in success or failure and discuss how you might quickly establish which of two different designs might be the more successful solution.

ii) Discuss with reasons which of the following problems are appropriate or inappropriate for solution by a genetic algorithm.

A. Maximising a unimodal differentiable function of several real variables over a closed bounded region.

B. Minimising a multimodal function of several continuous and discrete variables over a closed bounded region.

C. Maximising a convex objective function subject to linear constraints.

2. a Describe the Wisard model for pattern recognition. State some advantages of this model over other methods of pattern recognition.

b Suppose an N-tuple has been trained on patterns T_1, T_2. Let U be an unknown test pattern. Assume the retina area is normalised to unity, that the N-tuple inputs are randomly selected, and define

$$p_1 = |U \cap T_1|^N,\ p_2 = |U \cap T_2|^N,\ p_{12} = |U \cap T_1 \cap T_2|^N,$$

where $|S|$ denotes the area of a subset S of the retina.

Derive a formula for the probability that the N-tuple will fire when presented with the test pattern U.

3. A feedforward network has no hidden layers. If $x_1, \ldots, x_n$ are the inputs, the activation of the j *th* ($1 \le j \le m$) node of the output layer is defined as

$$net_j(x_1, \ldots, x_n) = \frac{1}{2\sigma_j^2} \sum_{r=1}^{n} (x_r - c_{jr})^2 \qquad (1)$$

where $c_{j1}, c_{j2}, \ldots c_{jn}$ and σ_j are adjustable parameters associated with the j *th* node. The node then computes the output

$$z_j = f(net_j) = \frac{1}{\sqrt{2\pi}} \exp\left(-net_j\right) \qquad (2)$$

The measure of error for the output vector is defined as

$$E(z_1, \ldots, z_m, t_1, \ldots, t_m) = \frac{1}{2}\sum_{j=1}^{m} \left(z_j - t_j\right)^2 \qquad (3)$$

where $(t_1, \ldots, t_m)$ are the target outputs for the inputs $(x_1, \ldots, x_n)$.

Show that a gradient descent learning rule to reduce E is given by

$$\Delta c_{jk} = \eta \delta_j \frac{1}{\sigma_j^2} (x_k - c_{jk}) \qquad (4)$$

and

$$\Delta \sigma_j = - \eta \delta_j \frac{1}{\sigma_j^3} \sum_{r=1}^{n} (x_r - c_{jr})^2 \qquad (5)$$

where

$$\delta_j = \frac{1}{\sqrt{2\pi}} \left(\exp(- net_j)\right) (z_j - t_j) \qquad (6)$$

Can this learning rule be extended to multilayer networks of such nodes? (You are **not** asked to attempt such an extension.) Briefly discuss the interpretation of such a model in terms of locally receptive fields of sensor cells.

Turn over...

4. a i) Define the *energy* E of a neural system with symmetric weights in which outputs of nodes are zero or one.

ii) Explain the operation of a stochastic unit as used in a Boltzmann Machine, and show how the probability of a node firing is related to energy.

iii) Sketch a graph of this probability against ΔE for different temperatures.

b As a simple demonstrator it is proposed to minimise the function

$$x_1x_2 - 2x_1x_3 + 5x_1x_4 - 3x_2x_4 + 3x_2x_3 - 7x_3x_4 + x_1 + x_2 + x_3 + x_4 \qquad (7)$$

over integer variables $x_i = 0$ or 1 $(1 \leq i \leq 4)$ using a Boltzmann machine.

i) Write down an appropriate matrix for the weights w_{ij} and a list of appropriate thresholds.

ii) Give a suitable algorithm for approximating the solution.

End of paper