IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2013

BEng Honours Degree in Computing Part III
BSc Honours Degree in Mathematics and Computer Science Part III
MSci Honours Degree in Mathematics and Computer Science Part III
MSc in Advanced Computing
MSc in Computing Science (Specialist)
for Internal Students of the Imperial College of Science, Technology and Medicine

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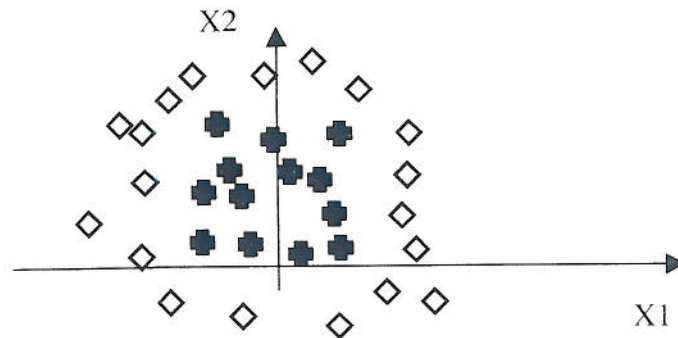*Associateship of the Royal College of Science*

PAPER C395

MACHINE LEARNING

Monday 13 May 2013, 14:30
Duration: 90 minutes

Answer TWO questions

1a Suppose that we want to build a neural network that classifies two dimensional data (i.e., X = |x1, x2|) into two classes: diamonds and crosses. We have a set of training data that is plotted as follows:

Draw a network that can solve this classification problem. Justify your choice of the number of nodes and the architecture. Draw the decision boundary that your network can find on the diagram.

1b Assume a learning rate η and a sigmoidal activation function in the layers of the network defined in 1a. What are the update rules for the weights of the output layer of the network? Compute (back-propagate) the errors of the hidden layer.

1c Derive the gradient decent training rule for a simple network with a single unit with output o given by the formula:

$$o = W0 + W1X1 + W1X1^2 + \dots WnXn + WnXn^2$$

where W are the weights of the network and X is the input. In your answer define explicitly the cost/error function E. You may assume, that D training examples with desired outputs (t_i, t_2, ..., t_D) are given.

1d What is the difference between the Best First Search and the Beam Search algorithms?

The four parts carry, respectively, 20%, 40%, 20%, 20% of the marks.

2a Suppose that we want to solve the problem of finding out what kind of music Alex likes by using genetic algorithms. Suppose further that the solution to the problem can be represented by a decision tree as follows:
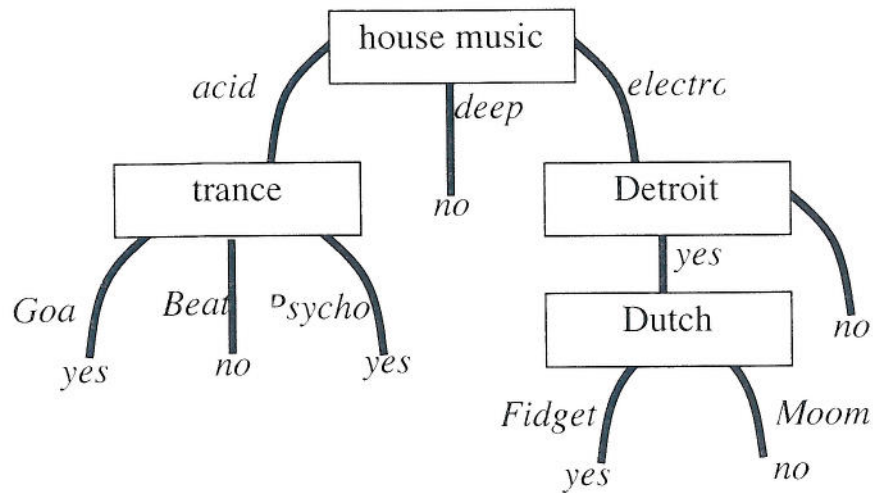
What is a suitable chromosome design for the given problem? Provide a short explanation of the solution.

What is the result of applying a single round of the prototypical Genetic Algorithm? Explain your answer in a clear and compact manner by providing the pseudo code of the algorithm. Provide a short explanation for each parameter used.

2b Consider the following set of positive (+) and negative (-) training examples:

house	trance	detroit	dutch	Alex likes
acid	goa	no	moom	+
electro	goa	yes	fidget	+
acid	beat	yes	fidget	-
deep	psycho	yes	moom	-
acid	goa	no	fidget	+

Apply a k-Nearest Neighbour algorithm, k=2, to classify the instance <electro, goa, no, fidget>, assuming that the above-listed examples are already known. Write out the algorithm, the distance function, and the intermediate results.

2c Which types of knowledge can we distinguish in Case-Based Reasoning (CBR)? Provide a short explanation of each of the types.

The three parts carry, respectively, 50%, 30%, 20% of the marks.

3a   Let

B_1 be $animal(X) \leftarrow fish(X)$

B_2 be $fish(wanda) \leftarrow$

$B = B_1 \wedge B_2$ be the background knowledge

$E = nice(wanda) \leftarrow$ be an example.

    i) Letting H stand for the hypothesis, state the condition in Inductive Logic Programming which H, B and E must satisfy.

    ii) For B and E above, what is $\perp$ (the most specific hypothesis)? Explain.

b   Let

a_1 be $p(3,4)$,

a_2 be $p(3,X)$,

a_3 be $p(Y,4)$ and

a_4 be $p(Z,Z)$.

    i) State the definition of $\succeq$ (subsumption) with respect to atoms.

    ii) For the atoms $a_1 - a_4$ above, state all true relations of the form $a_i \succeq a_j$.

    iii) For each pair of atoms $\langle a_i, a_j \rangle$ for which $i < j$ state the least general generalisation of a_i and a_j.

c   Consider the following two statements.

C: Francois lives in Paris
D: Francois lives in France

In each case below explain your answer.

    i) Represent statement **C** as a definite clause.

    ii) Represent statement **D** as a definite clause.

    iii) Define a background knowledge clause **B** which allows the clauses for **C** and **D** to be related according to their generality.

    iv) What is the generality relation between the clauses for **C** and **D** given **B**?

The three parts carry, respectively, 20%, 40%, and 40% of the marks.