IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2012

BEng Honours Degree in Computing Part III
MEng Honours Degree in Information Systems Engineering Part IV
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
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

Thursday 3 May 2012, 14:30
Duration: 90 minutes

Answer THREE questions

Paper contains 4 questions
Calculators required

Section A *(Use a separate answer book for this Section)*

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

	sky	air	humid	wind	water	forecast	Enjoy Sport
1	sunny	warm	normal	strong	warm	same	+
2	sunny	warm	high	strong	warm	same	+
3	rainy	cold	high	strong	warm	change	-
4	sunny	warm	high	strong	cool	change	+
5	sunny	warm	normal	weak	warm	same	-

1a Apply the CANDIDATE-ELIMINATION learning algorithm. Write out the intermediate and the final results.

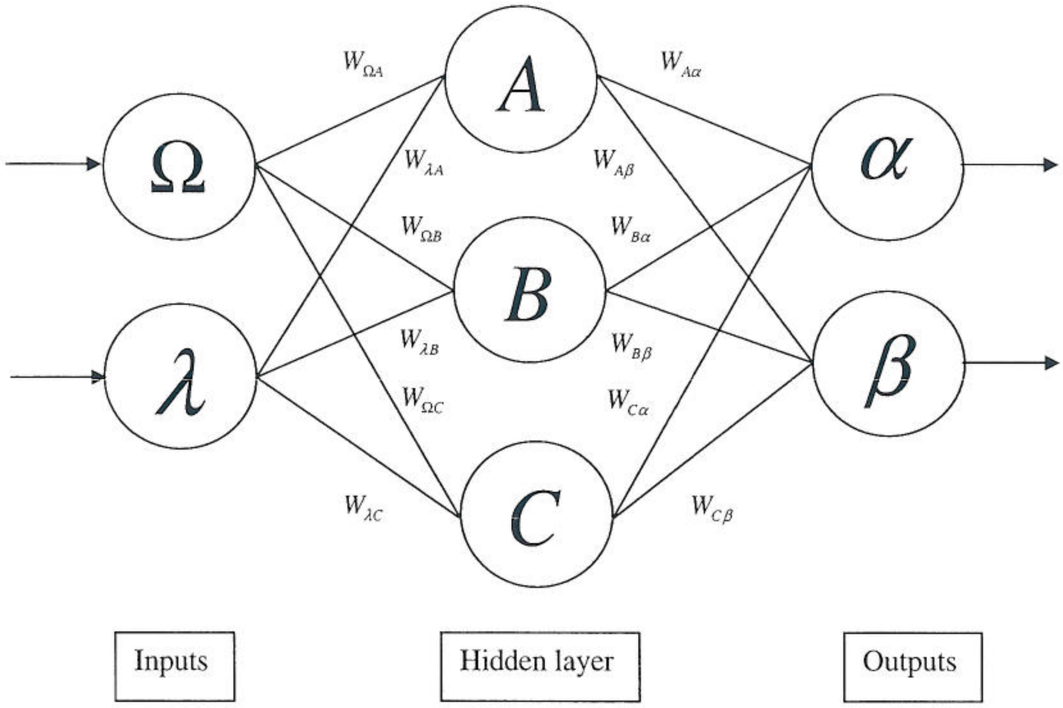
1b 1c Show the decision tree that would be learned by ID3 assuming that it is given the above-listed training examples. Write out the intermediate calculations.

1c What is the difference between the CANDIDATE-ELIMINATION and ID3 algorithms? What is the relationship between the learned decision tree in (1c) and the version space that is learned from the same examples in (1a)?

1d Apply a k-Nearest Neighbour algorithm, k=2, to classify the instance <sunny, warm, high, strong, cool, same>, assuming that the above-listed examples are already known. Write out the algorithm, the distance function, and the intermediate results.

The four parts carry, respectively, 30%, 30%, 10%, 30% of the marks.

2 Consider the 2-layers neural network given below. Assume a learning rate η and a sigmoidal activation function in both layers.



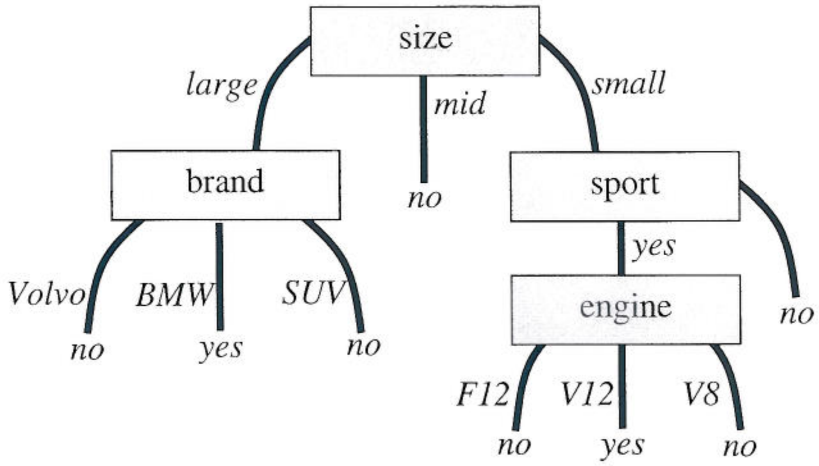
2a Compute the errors of the output layer. Write out the intermediate and the final results.

2b What are the update rules for the weights of the inner layer and the output layer?

2c Compute (back-propagate) the errors of the hidden layer. Write out the update rules for the inner layer weights.

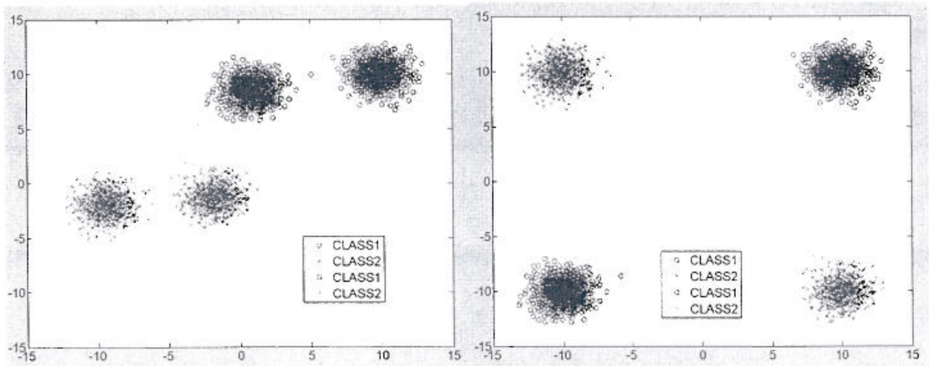
The three parts carry, respectively, 15%, 35%, 50% of the marks.

3a Suppose that we want to solve the problem of finding out what a good car is 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.

3b



Given the red and blue classes shown in the two figures above, explain what kind of Neural Networks could be used for classifying the samples in each case. Provide a short explanation of the suggested solutions.

3c Given a set of observations $D = \{(x_1, y_1), \ldots, (x_n, y_n)\}$, lets assume that the model $y_i = f(x_i) + e_i$ e_i follows $N(0, \sigma)$ and is independently drawn.
By assuming further mutually independent samples given f, prove that the Maximum Likelihood (ML) estimation for f^* is given by:

$$f^* = \text{argmin}_f \sum_{i=1}^{n} \left(f(x_i) - y_i\right)^2 .$$

3d Compute the ML estimation of λ in the case of $X_1, X_2, ..X_n$ being *iid* (independent and identically distributed) Poisson random variables (i.e, $P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$).

The four parts carry, respectively, 50%, 10%, 20%, 20% of the marks.