

**THE BCS PROFESSIONAL EXAMINATION
Professional Graduate Diploma**

April 2002

EXAMINERS' REPORT

Knowledge based Systems

Question 1

- a) Explain the principles of Case Based Reasoning (CBR) and describe the main techniques involved. (12 marks)
- b) Compare CBR with neural networks and rule based expert systems. (13 marks)

Answer Pointers

- a) A CBR system solves new problems by adapting solutions that were used to solve old problems.
- Retrieve the most similar cases
 - Reuse the cases to attempt to solve the problem
 - Revise the proposed solution if necessary
 - Retain the new solution as part of a new case.

A new problem is matched against cases in the case base, and one or more similar cases are retrieved. A solution suggested by the matching cases is then reused and tested for success. Unless the retrieved case is a close match, the solution will probably have to be revised, producing a new case that can be retained.

Techniques required are

- Case representation
- Indexing
- Storage
- Retrieval: Near neighbour retrieval, inductive retrieval
- Adaptation

- b) CBR vs Rule Based Expert Systems
- a reduction in knowledge elicitation effort
 - the ability to learn by acquiring new cases over time without adding new rules or modifying existing ones
 - the ability to provide justification by offering past cases as precedence rather than justifying a solution by showing a trace of the rules that led to the decision

CBR vs Neural Networks

- Some similarities
- Both rely on past cases
- NN is good in domains where data cannot be represented symbolically such as voice recognition
- CBR is less good with pure numeric data and much better with complex, structured symbolic data

(From Applying Case Based Reasoning by Ian Watson)

Question 2

- a) Explain why uncertainty management is necessary in expert systems. (9 marks)
- b) Use diagrams and examples to explain the concept of fuzzy sets and how associated operations work. (16 marks)

Answer Pointers

- a) Uncertainty management is necessary because of
- lack of data
 - inconsistency of data
 - imprecision in measurement
 - imprecision in concept
 - lack of a theory
- b) A set is defined by its members. A fuzzy set can be regarded as the label applied to a linguistically expressed concept which has no precise boundary. A fuzzy set can be represented numerically by a truth diagram where the y-axis is scaled from 0 (meaning false) to 1 (meaning true). Truth diagrams are expected in the answer.

It is possible that an object lies in the intersection of more than one set.

Operations like conjunction, disjunction, complementation and negation are required. It also required implication.

Examples.

(From Expert Systems by Richard Forsyth)

Question 3

- a) If an organisation had a knowledge management strategy, what would that mean? (10 marks)
- b) Describe, with examples, the process such an organisation would go through to elicit business knowledge from key staff. (15 marks)

Answer Pointers

- a) Knowledge management involves identifying and mapping intellectual assets within the organisation, generating new knowledge for competitive advantage and the enabling technology. The technology may include databases, knowledge based systems, knowledge engineers and data mining technologies.
- b) The process an organisation would go through to elicit business knowledge is called knowledge acquisition. This would include various interviewing and modelling techniques to elicit information from an individual. The results of which would be placed in a database / KBS.

Question 4

- a) Data mining is the process of extracting patterns and relationships from data. Describe two AI based algorithms that would commonly be used as part of a data mining process. (15 marks)
- b) It is said that 70% of the effort of a data mining project does not involve modelling at all. What activities would take place during this time? (10 marks)

Answer Pointers

- a) Some data mining algorithms:

CART: Classification And Regression Trees. CART is a method of splitting the independent variables into small groups and fitting a constant function to the small data sets. In categorical trees, the constant function is one that takes on a finite small set of values (e.g., Y or N, low or medium or high). In regression trees, the mean value of the response is fit to small, connected data sets.

GENETIC ALGORITHMS: A computer-based method of generating and testing combinations of possible input parameters to find the optimal output. It uses processes based on natural evolution concepts such as genetic combination, mutation and natural selection.

NEURAL NETWORKS: A complex non-linear modelling technique based on a model of a human neuron. A neural net is used to predict outputs (dependent variables) from a set of inputs (independent variables) by taking linear combinations of the inputs and then making non-linear transformations of the linear combinations using an activation function. It can be shown theoretically that such combinations and transformations can approximate virtually any type of response function. Thus, neural nets use large numbers of parameters to approximate any model. Neural nets are often applied to predict future outcome based on prior experience. For example, a neural net application could be used to predict who will respond to a direct mailing.

(Definitions for the answer were obtained from: <http://www.twocrows.com/glossary.htm>)

- b) 70% of the data mining process involves getting the data into the correct form ready for modelling. The data may be stored in different parts of the organisation, in different databases and in formats that are not useable by modelling techniques – whether it is because of format or level of aggregation.

Question 5

- a) Why do search techniques form the basis of many AI applications? (10 marks)
- b) For each of the following search techniques, describe how the algorithm works and provide an example of how it would be used.
- i) Breadth first
 - ii) Depth first (15 marks)

Answer Pointers

- a) Search techniques are critical to AI applications because – even with today's very fast computers – it is impossible to investigate all possible options within a given search space. This is because of the inherent time complexities of such problems.
- b) A description of how the algorithms work can be found in most good AI books.

Examples of where search techniques are used include computer chess and travelling salesman type problems.

Examiner's Comments

Part (a) was answered very poorly and part (b) was answered very well by most candidates. This demonstrated that candidates were well aware of the theory (as discussed in part (b)), but have very little understanding of where such theory could be used effectively, part (a).