THE BCS PROFESSIONAL EXAMINATION
Professional Graduate Diploma

April 2005

EXAMINERS' REPORT

Knowledge based Systems

Question 1

1. Knowledge based Systems (KBSs) are developed to deal with particular application domains in which alternative techniques are unable to produce reliable and manageable solutions.

a) Identify and discuss five aspects of human intelligence that could be used to characterise intelligent knowledge-based systems. **(10 marks)**

Answer Pointers

Question 1.a: 10 marks. 2 marks for each aspect of human intelligence identified and related to KBS.
Some possibilities include:

- Ability to temporarily alter behaviour according to environmental stimulus (**adaptability**)
- Ability to permanently alter behaviour as a result of accumulated experience with environmental stimuli (**learning**)
- Ability to deal with ill-defined and ambiguous situations (**uncertainty**)
- Ability to prioritise and focus (**goal directedness**)
- Ability to bring to bear subjective insight (**judgement**)
- Ability to deal with complexity and recognise relevance (**abstraction**)

Application of these traits to a medical diagnostic KBS should consider the points outlined below.

Adaptability: an investigative procedure can be refined as new symptoms emerge and new test results become available.

Abstraction: the most significant symptoms will be considered, and those may be considered as special cases of more general categories of symptoms. Connections between symptoms and causes may be established at the general level, e.g. fever is one of the flu-like symptoms, and flu-like symptoms could point to influenza, pneumonia, viral infection, etc.

Goal directedness: the diagnostic system has a main goal: to identify the cause of the symptoms; and will plan procedures that progress towards achieving that goal.

Judgement: recommendations and decisions will be based on experience with previous patient cases; thus, different experts may have different judgement (c.f. *second opinion*).

Learning: procedures for selecting and conducting tests are based on treatment of previous patients (with similar symptoms). Results of the treatment fed back to the KBS will be used to influence future decisions.

Uncertainty: prognoses about illnesses are often predictions that are based on incomplete information about the progression of the illness in the patient. That is, there can be incomplete information about the disease and about the patient.

b) Consider the requirements of different tasks, and specify features that warrant the application of knowledge-based technologies in only some tasks. Illustrate the discussion with one example of a task that demonstrates the need for a KBS, and one task for which KBSs are unnecessary. **(10 marks)**

Answer Pointers

Question 1.b: 10 marks. 2 marks for each feature of tasks identified and discussed in the context of KBS. 2 marks for each illustrative example.
Some possibilities include:

One way to characterise appropriate tasks is to test the task for its demands on the various characteristics of human intelligence described in Q1.a. We may summarise by noting that such tasks are *knowledge intensive*, which means that methods gained through experience and learning are needed in order to produce a solution. Such *knowledge-based* methods should be efficient when compared to the basic trial and error or brute force approaches in which every possibility is attempted until an acceptable solution is found (if it exists). The latter forms of method do not constitute intelligent behaviour.

A suitable task is one that can be tackled in many different ways, none of which can guarantee a satisfactory solution (but still offer the possibility of a solution) since this kind of task would require some of the characteristics introduced earlier (e.g. judgement and reasoning with uncertainty). Architectural design is a good example of this kind of *open-world* or *ill-defined* problem, which needs *approximate* or *heuristic* reasoning.

Examples of knowledge intensive tasks are:

- Diagnosis of malfunctions or diseases
- Configuration of complex objects such as computer systems or designing offices
- Planning sequences of actions such as scheduling timetables

In contrast, a problem that may be solved by following a step by step procedure (e.g. a formula), which is well determined and guaranteed to produce a solution would not qualify as demanding intelligence: e.g. solving simple algebraic problems in maths, or determining standard statistical procedures. These kinds of problem may be solved *algorithmically* by any computing machine that can follow a procedure.

It should be noted that any task can be tackled using intelligent methods, though, unless there is a significant improvement in efficiency from using intelligent methods, the task should not be admitted as one that demands intelligence.

Furthermore, many tasks can be reduced to the problem of searching for a solution in a search space. Accordingly, much of artificial intelligence work is involved with finding effective ways of structuring and representing the search space, and with generating efficient methods for conducting searches. Tasks that push the limits in either of these areas are good candidates for intelligent KB systems.

c) Comment briefly on the reasons for an apparent disproportionately low use of KBSs compared to conventional systems in solving business problems. **(5 marks)**

Answer Pointers
Question 1.c: 5 marks. General distribution of marks according to salient points.
Some points to note include:

Perceptions of the reliability, complexity, cost and availability of suitable KBS for solving practical business problems.

Suitability of business problems for solution through KB methods – see Q1.b.

Cost and time required to develop KB systems.

Availability of qualified knowledge engineers to develop bespoke systems, in contrast to off-the-shelf non-intelligent software.

Examiner's Comments
The majority of students attempted this question.

Question 2

2. Knowledge elicitation involves modelling the knowledge used by an expert to solve problems. Consider an example application domain (different from that of Question 3) and construct a knowledge base for the domain by completing the following tasks:

a) Describe briefly general methods that could be applied to elicit the knowledge needed to solve a small complex problem. **(5 marks)**

b) Identify a suitable task, and produce a decision tree of five levels to represent the decision process. **(5 marks)**

c) Illustrate the consistency of the decision tree with a worked example using sample data for the task (specify any assumptions made). **(5 marks)**

d) Describe how an implemented knowledge base could be tested. **(5 marks)**

e) Comment on how machine learning could be utilised to make subsequent decision making with the
knowledge base more efficient. **(5 marks)**

Answer Pointers
Question 2: General distribution of marks for each subsection according to salient features being present.
Q2.a:
Techniques such as interviewing, observation, and automated elicitation systems should be discussed.

Q2.b:
Any task that affords five levels of decision tree should be stated, e.g. diagnosing faults in a photocopier. Deduct 3 marks if the same domain as in Q3 is chosen. A sample decision tree for deciding on the amount of discount to give in a transaction is given below for three levels.

Q2.c:
Assume the following scenario: cash sale of £75.

The decision process would flow through the lower branch at level 1, then the centre branch (Sales=50-100), to conclude that a 10% discount is warranted.

The examinee must work through five levels.

Q2.d:
Testing of the knowledge based system should involve systems evaluation, i.e. verification and validation that it satisfies the requirements an design. Also, it needs to be tested to ensure that its solutions are reliable. This can be done by comparing a sample of solutions prepared by with a group of experts.

Q2.e:
Techniques such as case-based reasoning or rule-induction, etc. could be used to maintain the knowledge base over time as it gains experience in solving problems. For non symbolic systems, continuous training and evaluation techniques could be employed to evolve the system.

Examiner's Comments
A fairly popular question.

Question 3

3. Consider the following logical argument.

"Every KBS is intelligent; CYC is a KBS, therefore, CYC is intelligent".

a) Explain why the argument could not be validly symbolised using propositional logic and present the predicate logic form for the argument. **(5 marks)**

b) Construct a rule model for deciding on the choice of a new car. Include the following features:

i) at least ten rules (no more than 15) represented in both natural language and predicate logic form, which must be capable of being data driven to make a decision, with at least five assertions made.
(You should make use of quantifiers, binary relations, and Boolean operators). **(15 marks)**

ii) a worked example showing the decision making process with sample data. **(5 marks)**

Answer Pointers
Question 3.a: 3 marks for discussion and 2 marks for predicate logic form of argument.
Q3.a:
The argument is read literally as "For all objects in the universe of discourse, if an object is a KBS, then that object is intelligent. CYC is a KBS, therefore, CYC is intelligent."

This is a valid argument because, by virtue of the universal quantifier (for all), we have introduced the conditional to make a connection between "KBS" and "intelligent."

Once we establish that "CYC" qualifies to be considered a KBS, we simply apply the conditional to deduce that CYC is intelligent.

The key step has been to recognize that words such as "every", "all" and "any" implicitly include any particular example of that thing. For instance, "any car is bad for the environment" refers to every car that could be selected: yours, mine, the Queen's, etc. Therefore, if we form any proposition from such words we automatically allow any particular example of that thing to be included in the proposition. "Any car is bad for the environment" justifies the acceptance of the following propositions (plus many other similar ones):
My car is bad for the environment.
Your car is bad for the environment.
The Queen's car is bad for the environment.

Propositional logic does not cater for quantifiers such as every, and all, hence it is limited in its expressive power.

So if we have an argument:

Every KBS is intelligent; CYC is a KBS, therefore, CYC is intelligent.

Using propositional logic we *could not* form a valid argument.

P Every KBS is intelligent
Q CYC is a KBS
R CYC is intelligent

P, Q, $\vdash$ R Fallacy.
But using predicate logic and the universal quantifier we *could*.

P(x) x intelligent (predicate)
Q(x) x is a KBS (predicate)
C CYC (individual constant)
x (individual variable)

$(\forall x)(Q(x) \rightarrow P(x)), Q(x) \vdash P(x)$ valid.

b) Construct a rule model for deciding on the choice of a new car. Include the following features:

ii) at least ten rules (no more than 15) represented in both natural language and predicate logic form, which must be capable of being data driven to make a decision, with at least five assertions made. (You should make use of quantifiers, binary relations, and Boolean operators). **(15 marks)**

ii) a worked example showing the decision making process with sample data. **(5 marks)**

Answer Pointers

Question 3.b.i: 3 marks for each rule up to 10 rules maximum – 1 mark for natural language version and 2 marks for predicate logic version. Deduct 1 mark for each required feature not used up to 5 marks in total.

Q3.b.i:

A set of cohesive rules stated in predicate logic are needed. They should all **follow the general patterns given below which use the conditional.**

$\forall(x)(P(x) \rightarrow Q(x)),$

$\forall(x)\ \forall(y)(Q(x) \rightarrow Q(y)),$

$\forall(x)\ \forall(y)\ \forall(z)\ ((Q(y) \wedge P(z)\ ) \rightarrow R(x,y,z))$

These rules should be chained.

Question 3.b.ii: 5 marks distributed for reasoning chain.

Q3.b.ii:

From the set of rules, a series of valid logical arguments must be presented using the predicate form of rules.

$\forall(x)(P(x) \rightarrow Q(x)),$

$\forall(x)\ \forall(y)(Q(x) \rightarrow Q(y)),$

$\forall(x)\ \forall(y)\ \forall(z)\ ((Q(y) \wedge P(z)\ ) \rightarrow R(x,y,z))$

These rules should be chained, e.g.

$(\forall x)(P(x)\rightarrow Q(x)),\ P(x)\ \vdash\ Q(x)$ **valid.**

$(\forall x)\ (\forall y)(Q(x)\rightarrow Q(y)),\ Q(x)\ \vdash\ Q(y)$ **valid.**

$(\forall x)\ (\forall y)\ (\forall z)((Q(y) \wedge P(z)\rightarrow R(x,y,z)),\ Q(y),\ P(z)\ \vdash\ R(x,y,z)$ **valid.**

Hence, from input data P(x) and P(z), the rules have been chained to derive R(x,y,z).

Question 4

4. Describe how the following data mining operations are applied and provide typical examples for each:

i) Predictive modelling **(6 marks)**
ii) Database segmentation **(6 marks)**
iii) Link analysis **(7 marks)**
iv) Deviation detection **(6 marks)**

Answer Pointers

Each sub question is marked 6.5 marks:

- Predictive modelling can be used to analyse an existing database to determine some essential characteristics (model) about the data set. The model is developed using a supervised learning approach, which has two phases: training and testing. Predictive modelling could be applied for customer retention management, credit approval, cross-selling direct marketing.
- Database segmentation: partition a database into an unknown number of segments or clusters of similar records. This approach uses unsupervised learning to discover homogeneous sub-population in a database to improve the accuracy of the profiles.
- Links analysis aims to establish links, called association. There three specialisation of link analysis: association discovery for finding items that imply the presence of other items in the same event, sequential patterns discovery for finding patterns between events such that the presence of a set of items is followed by another set of items in a database of events over a period of time, similar time sequence discovery is used, for example, in the discovery of links between two sets of data that are time-dependent, and is based on the degree of similarity between the patterns that both time series demonstrate.
- Deviation detection is often a source of true discovery because it identifies outliers, which express deviation from some previously known expectation and norm. This operation can be performed using statistics and visualisation techniques or as a by-product of data mining.

Question 5

5. A bus company has to set the optimal route for its bus, which has to go through four cities and return to its city of origin. Each city has to be visited exactly once, and the cost of the journey between each pair of cities is known. The problem is to do the tour at minimum total cost or distance.

From/To	**A**	**B**	**C**	**D**
A	-	4	6	12
B	3	-	6	8
C	7	10	-	10
D	11	7	9	-

Table 1: Distance (or cost) between Cities

a) Illustrate your description by explicitly finding the shortest tour given the distances (or costs) between cities shown in **Table 1** above. **(8 marks)**

b) Describe briefly how a solution to the Bus Journey Problem might be obtained using a genetic algorithm, indicating how features of the problem map to the elements needed to use a genetic algorithm. **(12 marks)**

c) Describe briefly how a solution to the Bus Journey Problem might be obtained using simulated annealing, again indicating how you would obtain the necessary features for *this* problem. **(5 marks)**

Answer Pointers

a) We show the calculation below. We have chosen *A* as the starting point and listed all possible tours; there is a unique tour with the minimum length of 26, namely *ACDB*.

Tour	Step 1	Step 2	Step 3	Step 4	Total
A B C D	4	6	10	11	31
A B D C	4	8	9	7	28
A C B D	6	10	8	11	35
A C D B	6	10	7	3	26
A D B C	12	7	6	7	32
A D C B	12	9	10	3	34

b) The two essential ingredients of a genetic algorithm are the availability of a natural representation of the problem and its constraints, which leads to suitable genetic operations, and the availability of a useful fitness function. In this problem, the fitness function is clear; two tours can be evaluated simply by looking at the total cost of each tour; in contrast there is no natural representation, which describes the problem.

Of course a naive binary representation of each tour is available, in which each city is numbered, and a gene is simply an ordered list of city numbers. This representation does not however lend itself to useful genetic operations; a naive mating may not even produce a tour (some city numbers may be invalid), and is unlikely to be a valid tour with no repeat cities.

We are thus led to seek alternative (non-binary) representations, which are adapted to this particular problem. Here are some possible representations.

- The ordinal representation, in which the tour (1, 2, 4, 3, 8, 5, 9, 6, 7) is represented as the list (1, 1, 2, 1, 4, 1, 3, 1, 1) in which the code refers to the position of the next city of the list of unused cities.
- The path representation, which was the natural one we used above; and
- The adjacency representations in which the tour (1, 2, 4, 3, 8, 5, 9, 6, 7) in the path representation is given as (2, 4, 8, 3, 9, 7, 1, 5, 6), indicating by the entry in position *k* that the tour goes from *k* to this new city.

Even when crossover is defined, it does not lead to natural ``genetic evolution'' and so it is necessary to have much more specialised crossover operators, in which the naive crossover is ``repaired''.

c) We simply use the obvious path representation described above; thus each potential tour can be described by a re-arrangement **i** of the city labels.

Examiner's Comments
This question assesses the students understanding on how genetic algorithm is working. Students are advised to use the provided Bus Journey Problem to demonstrate their understanding. Most of the students have answered correctly the sub-question (a) by finding the shortest tour given the distances (or costs) between cities. However, less than a half of students have responded correctly to sub questions (b) and (c) in which they are asked to describe a solution to the Bus Journey Problem using a genetic algorithm and simulated annealing techniques. This question is useful for those students with less expressive power in their narrative to gain marks by answering practical questions. There is a need for student to be given more opportunity to practice problem solving using genetic algorithm.