THE BCS PROFESSIONAL EXAMINATION
BCS Level 6 Professional Graduate Diploma in IT

April 2008

EXAMINERS' REPORT

Knowledge Based Systems

General Comments

Table 1 shows the statistics based on all questions answered, including where some students answered four questions. Averages are fairly consistent across questions, though, a little on the low side. Standard deviation results are fairly low across all questions. Thus, it may be concluded that the candidates appear to be fairly equated in their ability, but perhaps were overly challenged by the examination (in contrast to the previous year). Questions 1 and 5 were the most popular amongst candidates, whereas the more technology oriented questions (3 and 4) were attempted by the fewest candidates, possibly due to the discurive nature (as was noted for the previous year). It appears that candidates prefer specific questions on KBS issues and technologies rather than general discursive type questions on broader issues (again).

TABLE 1: statistics for all questions answered

	Q1	Q2	Q3	Q4	Q5	Total
Examiner (initials)						
Number Attempted	43	26	21	16	45	
% Attempted	89.58%	54.17%	43.75%	33.33%	93.75%	
Number Accepted	42	24	19	15	44	
% Accepted	87.50%	50.00%	39.58%	31.25%	91.67%	
Number Passed	40	20	15	10	42	45
% Passed	95.24%	83.33%	78.95%	66.67%	95.45%	93.75%
Max Mark	19	18	20	15	21	74.67
Min Mark	6	5	1	4	4	22.67
Average Mark	13.76	11.33	11.79	10.53	15.07	52.67
Standard Deviation	2.93	3.00	4.25	3.24	3.61	9.72

Seven students answered four questions instead of three. When the top three questions are considered alone, the average is 52.5 and standard deviation is 9.78.

Question 1

Development of any industrial level software system requires careful management in order to ensure that an effective product is produced. Describe a project management framework that is suitable for the construction of an interactive web based intelligent system by addressing the following tasks:

i. Identify all the main stakeholders involved in the project, and explain their respective roles and responsibilities in the construction of an intelligent knowledge based system. **(10 marks)**

ii. Identify a suitable project management and software development methodology, and describe the main phases in the development process as they relate to knowledge engineering in particular. **(15 marks)**

Question 1.i: 10 marks. General distribution of marks according to salient features.

There are generally three individuals having an interaction with expert systems. Primary among these is the end-user; the individual who uses the system for its problem solving assistance. In the building and maintenance of the system there are two other roles: the problem domain expert who builds the knowledge base, and a knowledge engineer who assists the experts in determining the representation of their knowledge and who defines the inference technique required to obtain useful problem solving activity. Additionally, other stakeholders are the project manager, who takes operational control of the project, and the board of directors and clients, who are responsible for commissioning and overseeing the project.

Question 1.ii: 15 marks. General distribution of marks according to salient features.

Knowledge Acquisition and Documentation Structuring KADS is a structured way of developing knowledge-based systems (expert systems). It was developed at the University of Amsterdam as an alternative to an evolutionary approach and is now accepted as the European standard for knowledge based systems.

Its components are:

A methodology for managing knowledge engineering projects.
A knowledge engineering workbench.
A methodology for performing knowledge elicitation.
KADS was further developed into CommonKADS.

Examiners' Comments

A high proportion of the candidates attempted this question (85.42%) and on average had only fair results (13.95/25, standard deviation 2.70). Part 1 was answered reasonably due to the fact that there were many acceptable stakeholders who could be identified. However, many candidates did not offer sufficient discussion of each stakeholder's roles and respopnsibilities. Part 2 was not answered well because candidates failed to identify a specific project management methodology. Furthermore, most candidates discussed software engineering instead of knowledge engineering.

Question 2

Knowledge elicitation involves modelling the knowledge used by an expert to solve problems. Consider an example application domain and construct a knowledge base for the domain by completing the following tasks:

i. Describe briefly general methods that could be applied to elicit the knowledge needed to solve a small complex problem. **(5 marks)**
ii. Explain why it would be difficult to develop a KBS for a domain in which there was a considerable reliance on tacit and implicit knowledge. **(5 marks)**
iii. Present an example knowledge base using a knowledge representation formalism of your choice. Ensure that it is adequately annotated with a textual explanation of how it could be used to solve problems. **(15 marks)**

Answers Pointers

Question 2.i: 5 marks. General distribution of marks according to salient features.
Techniques such as interviewing, observation, and automated elicitation systems should be discussed.

Question 2.ii: 5 marks. General distribution of marks according to salient features.
The nature of tacit and implicit knowledge is such that describing them in a tangible form is difficult. Such knowledge is demonstrated in use rather than being amenable to representation. KBS that attempt to emulate tacit and implicit knowledge need to overcome the problem of necessarily depending on an executable model of knowledge. A solution is to develop methods that allow tacit and implicit knowledge to emerge during execution of explicit knowledge – very challenging though.

Question 2.iii: 15 marks. General distribution of marks according to salient features.
Candidates may choose any recognized knowledge representation formalism, e.g. rules, cases, frames, logic, etc. A clear explanation of the knowledge base must be provided from which it should be possible to understand how the KB would be utilised to solve problems.

Examiners' Comments

One half of candidates (50.00%) answered question 2 and made a fair attempt (average 11.33/25; standard deviation 3.00). Part 1, identification of elicitation methods, was straightforward, though often specialist KBS methods were not given enough attention. Part 2 was more problematic since candidates failed to understand the characteristics of tacit and implicit knowledge, and therefore, could not relate those characteristics to problems with knowledge elicitation and acquisition. Worked examples in part 3 often did not provide a trace of the decision-making path, and failed to show how the chosen problem would be solved.

Question 3.
Data mining is arguably one of the most successful applications of artificial intelligence in business.

i. Discuss how it has successfully made the transition from research laboratory to eBusiness applications. Focus on the real business problems that the technology addresses and the benefits perceived to have been realised. **(20 marks)**
ii. Comment briefly on the reasons for an apparent disproportionately low use of KBS compared to conventional systems in solving business problems. **(5 marks)**

Answer Pointers

Question 3.i: 20 marks. General distribution of marks according to salient features.
Data mining is often used in business, particularly on the web, to provide support for the "implict web", in which personal information about the user is indirectly "discovered" from their normal interactions. Such valuable knowledge enables tailoring and personalization of services.

Question 3.ii: 5 marks. General distribution of marks according to salient features.
AI solutions have specialized applications and tend to be costlier to produce. They are also less well understood, which makes them less reliable, and consequently they are more susceptible to hype. For instance, fuzzy logic is often used in consumer products such as cameras (autofocus), or washing machine controllers. Data mining is now often employed in business, intelligent search engines on the web (Google), games, consumer robots with vision systems, handwriting reading in PDA, speech synthesis in word processors, intelligent tutoring, knowledge management, etc. As knowledge working grows, it is expected that such technologies would continue to become more prevalent.

Examiners' Comments

About two fifths of candidates (41.67%) answered question 3 and made a poor attempt (average 11.79/25; standard deviation 4.25). For part 1, answers focussed on describing the technology rather that discussing its application to real problems. And for part 2, answers tended to be limited to cost factors and lack of development expertise and did not address the suitability of of each kind of software system to business needs.

Question 5

AI technologies each offer different advantages and limitations. Select two opposing technologies and undertake the following analysis:

i. Compare and contrast their features and indicate criteria that could be applied to help select the most suitable technology for a given application. **(15 marks)**

ii. Consider how the two chosen technologies could be combined to produce a hybrid system that would, in principle, overcome the limitations inherent in each individually. **(10 marks)**

Answer Pointers

Question 4: 25 marks. General distribution of marks according to salient features.

A wide range of technologies could be discussed. Comparing neural nets with rule based systems. An ability to provide explanations, or to deal with applications in which knowledge is implicit, an ability to learn and self-develop. Others include

Neuro-fuzzy systems, fuzzy expert systems, and connectionist expert systems.

Connectionist expert systems are artificial neural network (ANN) based expert systems where the ANN generates inferencing rules e.g., fuzzy-multi layer perceptron where linguistic and natural form of inputs are used. Apart from that, rough set theory may be used for encoding knowledge in the weights better and also genetic algorithms may be used to optimize the search solutions better.

Neuro-fuzzy expert systems are hybrids of artificial neural networks and fuzzy logic. Neuro-fuzzy hybridization results in a hybrid intelligent system that synergizes these two techniques by combining the human-like reasoning style of fuzzy systems with the learning and connectionist structure of neural networks. Neuro-fuzzy systems incorporate the human-like reasoning style of fuzzy systems through the use of fuzzy sets and a linguistic model consisting of a set of IF-THEN fuzzy rules. The main strength of neuro-fuzzy systems is that they are universal approximators with the ability to solicit interpretable IF-THEN rules.

Examiners' Comments

About one third of candidates (31.25%) answered question 4 and made a poor attempt (average 10.53/25; standard deviation 3.24). Answers tended to describe two or sometimes several AI technologies separately rather than providing a juxtaposition of two. Furthermore, answers failed to address the issues of hybridization, and instead, offered solutions based on two independent systems being applied to a common problem task.

Question 5

There are many alternative ways to solve a problem.

i. Explain both brute-force and heuristic search methods and discuss their relative merits with the aid of suitable examples. **(10 marks)**

ii. Explain the difference between inductive and deductive reasoning. **(5 marks)**

iii. With the aid of an illustrative example, describe how a problem could be solved using analogical reasoning. **(10 marks)**

Answer Pointers

Question 5.i: 10 marks. General distribution of marks according to salient features.

A searching algorithm requires a target for which to search. The list is searched until either the target is located or the algorithm has determined that the target is not in the list. A comparison must be made to determine if the current element retrieved from the collection is the target one; therefore, a measure of similarity is needed. One of the fields, called the *key* field, serves as the measure on which comparison is performed.

The set of all possible solutions to a problem is called the search space. Brute-force search or uninformed search algorithms use the simplest, most intuitive method of searching through the search space, whereas informed search algorithms use heuristics to apply knowledge about the structure of the search space to try to reduce the amount of time spent searching.

Uninformed search

An uninformed search algorithm is one that does not take into account the specific nature of the problem. As such, they can be implemented in general, and then the same implementation can be used in a wide range of problems thanks to abstraction. The drawback is that most search spaces are extremely large, and an uninformed search (especially of a tree) will take a reasonable amount of time only for small examples. As such, to speed up the process, sometimes only an informed search will do.

Informed search

In an informed search, a heuristic that is specific to the problem is used as a guide. A good heuristic will make an informed search dramatically out-perform any uninformed search.

There are few prominent informed list-search algorithms. A possible member of that category is a hash table with a hashing function that is a heuristic based on the problem at hand. Most informed search algorithms explore trees, such as the Best-first search, which is a search with a heuristic that attempts to predict how close the end of a path is to a solution, so that paths which are judged to be closer to a solution are extended first. Efficient selection of the current best candidate for extension is typically implemented using a priority queue.

Question 5.ii: 5 marks. General distribution of marks according to salient features.

Inductive reasoning is the process of reasoning in which the premises of an argument are believed to support the conclusion but do not ensure it. It is used to ascribe properties or relations to types based on tokens (i.e., on one or a small number of observations or experiences); or to formulate laws based on limited observations of recurring phenomenal patterns. Induction is employed, for example, in using specific propositions such as:

This ice is cold.
A billiard ball moves when struck with a cue.
...to infer general propositions such as:
All ice is cold.
All billiard balls struck with a cue move.

Deductive reasoning is the type of reasoning that proceeds from general principles or premises to derive particular information. It applies general principles to reach specific conclusions, whereas inductive reasoning examines specific information, perhaps many pieces of specific information, to derive a general principle. Deductive reasoning is supported by deductive logic.

For example:
All apples are fruit.
All fruits grow on trees.
Therefore all apples grow on trees.

Question 5.iii: 10 marks. General distribution of marks according to salient features.

Case-based reasoning (CBR), broadly construed, is the process of solving new problems based on the solutions of similar past problems. An auto mechanic who fixes an engine by recalling another car that exhibited similar symptoms is using case-based reasoning. A lawyer who advocates a particular outcome in a trial based on legal precedents or a judge who creates case law is using case-based reasoning. So, too, an engineer copying working elements of nature (practicing biomimicry), is treating nature as a database of solutions to problems. Case-based reasoning is a prominent kind of **analogy** making.

Examiners' Comments

Question 5 was the best answered (average 15.07/25; standard deviation 3.61), with the most candidates attempting it (91.67%). Parts 1 and 2 were straighforward and posed few problems for candidates, though answers varied in their level of precision and detail. For Part 1, provided examples were often inadequate, while for Part 2, some answers confused the difference between inductive and deductive reasoning. Part 3 was a bit more problematical since answers did not contain accurate explanation of analogical reasoning (as opposed to similarity based reasoning) and examples tended to be ineffective.