

**THE BCS PROFESSIONAL EXAMINATION
Professional Graduate Diploma**

April 2003

EXAMINERS' REPORT

Advanced Database Management Systems

Question 1

1. Describe the characteristics of the infrastructure for mobile computing, and explain how database systems operate within it. **(10 marks)**

Discuss some of the areas of difficulty in providing mobile database services and how these may be overcome. **(15 marks)**

Examiner's Comments

On the whole this question was answered reasonably well, although in some cases the descriptive information for the first part was sparse or superficial.

Answer Pointers

Mobile Computing Infrastructure :

Candidates may describe the following and their interrelationships

- Personal computers, Laptops, Notebooks
- Wireless Local-Area Networks
- Packet-based Cellular Telephony
- Mobile Hosts
- Mobile Support Stations
- Cell
- Mobile Host Movement between Cells
- Control Handoffs
- Direct Communication between hosts e.g. Bluetooth
- Wireless Application Protocol
- Wireless Markup Language

(10 marks)

Difficult Areas:

Routing and Query Processing

Mobility affects database query processing. There are dynamically changing communication costs. Optimisation becomes more complex.

Competing costs in optimisation of queries:

- User time
- Connection time
- Unit of transmission
- Time of Day Charges
- Energy

Broadcasting is a solution for frequently requested data.

Disconnectivity and Consistency

Mobile hosts to be disconnected for substantial periods. This may be because of cost. During periods of disconnection the mobile host may remain in operation with respect to local data. However updates to data may become inconsistent. This can be solved by version numbering. The version-vector scheme detects inconsistencies when copies of a document are independently updated. This scheme however does not address the reconciliation of inconsistent copies of data. It also needs substantial communication between a reconnecting mobile host and that host's mobile support station. The potential for disconnection and the cost of wireless communication limit the practicality of traditional transaction processing techniques. Long-term blocking can occur because of disconnection.

Security issues

Security is an important aspect to consider also.

Costs

Transmission costs may be an issue which leads to the disconnectivity and consistency problem described above.

(15 marks)

Further information about this can be found in the set text A SILBERSCHATZ, H KORTH and S SUDARSHAN, Database System Concepts, McGraw-Hill

Question 2

2. There are a number of algorithms for computing a join of two relations. Describe THREE such algorithms. (15 marks)

Using illustrations, compare and contrast them in terms of efficiency. (10 marks)

Examiner's Comments

The main problem in the answers to this question was that candidates described types of join (e.g. left join, right join, natural join, equi join) rather than the different sorts of underlying algorithm for a join.

Answer Pointers

There are a number of join methods that could be described:

Simple evaluation
Block-oriented evaluation
Hash Join
Sort Merge Join
Indexed Join

Candidates should describe the algorithm for three of these. They should then compare the efficiency by means of an illustration.

Marking Scheme

Description of algorithm	15 marks
Comparison and illustration	10 marks

Question 3

3. Explain why transaction management is an essential part of multi-user database systems. (10 marks)

Evaluate TWO common methods for ensuring integrity of data in concurrent transaction processing, with reference to the following schedule.

Transaction A	Transaction B
Read (A)	
	Read (A)
Read (B)	
	Read (B)
Write (B)	
	Write (A)
Write (A)	
	Write (B)

(15 marks)

Examiner's Comments

This question was answered reasonably well. Many candidates understood how locking is applied but were less sure about timestamping.

Answer Pointers

Transaction management is important so that the database is not left in an inconsistent state as a result of competing transactions. ACID properties may be discussed. Concurrency problems such as lost update, inconsistent analysis and uncommitted dependency may be described.

(10 marks)

The two methods of ensuring consistency are time-stamping and locking. Concept of serialisability may be discussed when describing these. The given schedule is not serialisable. Candidates should show how locking and timestamping would be applied to the given schedule. With locking transaction A fails when an exclusive lock is requested to write B as transaction B has already been granted a shared lock. In timestamping transaction A fails at the same place because transaction B which started later has already read B, an item that should have been updated by A before transaction B reads it (according to serialisability rules).

(15 marks)

Question 4

4. a) Explain what is meant by the term 'constraint database system'. (12 marks)

b) Compare the way spatial data is handled in:
i) constraint databases
ii) relational and object-oriented databases. (7 marks)

c) Compare the way temporal data is handled in:
i) constraint databases
ii) relational and object-oriented databases (6 marks)

Examiner's Comments

This was not a popular question and a number of the answers indicated that this part of the syllabus has not been covered well in the taught courses. Many candidates incorrectly attempted to describe constraints in the relational model. This was not required, the question was clearly concerned with constraint databases. Those who did understand this topic were often only able to answer the first part of the question and were not able to describe the use of constraint based systems for spatial and temporal data.

Answer Pointers

Constraint databases view data as constraints and constraints as data. Consider the relational scheme EMP(Empno, Ename, Sal, Deptno). One tuple in such a scheme might be (E1, Jones, 10000, D1). In a constraint database this could be represented as $\{(Empno, Ename, Sal, Deptno) \mid Empno=E1 \wedge Ename=Jones \wedge Sal=10000 \wedge Deptno=D1\}$. Here $Empno=E1$ is an atomic constraint and \wedge is the logical AND operator. Constraints may be constructed by logical operations on other constraints so therefore the tuple is defined by the single constraint given. A generalised tuple has the form $p_n(A_1, \dots, A_m) \mid C$ where C is a constraint on the attributes A_i and m is the arity of the relation. Relational databases are a special case of constraint databases where the atomic constraints are equations. In the more generalised case atomic constraints may contain inequalities and specify more than one tuple. Since constraints are logical expressions it is possible to use analytic techniques to optimise queries on such databases. (12 marks)

Relational databases are not ideal for expressing spatial data. Such data is normally stored as a set of polygons represented as a set of points. The number of sides of the polygon will vary. Whilst this can be stored in a relational database the processing of such data is normally inefficient. Object-oriented databases are more amenable to the storage of this type of data but lack indexing and optimisation techniques. Constraint databases can easily represent a polygon as a constraint and can optimise queries and define indices. (7 marks)

Temporal databases require a system to define infinitely large sets of points since time can be sliced into chunks as small as a user requires. Neither relational nor object-oriented databases are especially useful in this respect. Consider, however, the following $\{(Empno, Ename, Sal, Deptno) \mid Empno=E1 \wedge Ename=Jones \wedge Sal > 5000 \wedge Sal < 15000 \wedge Deptno=D1\}$. If Sal is a real number this expression defines an infinite number of tuples since there are an infinite number of real values between 5000 and 15000. This feature of constraint databases makes them particularly suited for representing temporal data. (6 marks)

Question 5

5. a) Describe the CORBA architecture. **(9 marks)**
- b) Discuss the reasons why combining a CORBA implementation with database technology may:
- i) benefit a CORBA application **(8 marks)**
 - ii) increase the functionality offered by a database management system. **(8 marks)**

Examiner's Comments

Not a popular question but in general answered rather better than question (4). Most candidates who attempted this question knew what CORBA is and were able to offer a good description of it. A smaller number were able to assess the interplay between CORBA and database management systems which was explored in the second two parts of the question.

Answer Pointers

CORBA (Common Object Request Broker Architecture) is an OMG standard. It allows access to object across a network independently of their location or the programming language used to construct them. Each object defines the services it offers using a standard interface definition language (IDL).

(9 marks)

From the point of view of a CORBA application, databases offer a mechanism for persistence. Complex operations will often require objects to store their data over long periods of time. CORBA provides a mechanism by which objects may make use of a database and it is therefore relatively easy to implement persistence via the database.

(8 marks)

From the database viewpoint CORBA may offer the following advantages:

- Multiplatform support: distribution, OS, programming language
- Integration between the DBMS and other components in the system
- Lightweight clients (distribution implemented in CORBA)

(8 marks)