THE BCS PROFESSIONAL EXAMINATION
Diploma

April 2000

EXAMINERS' REPORT

Database Systems

Question 1

Answer Pointers

(a) This part of the question first seeks a brief definition of what a foreign key is. Then the question seeks to determine the candidate's practical understanding of the role of foreign keys.

Marking scheme: two marks for a definition of the concept of a foreign key (one for basic concept and one for details); two marks to explain why foreign keys are important; two marks for an example of use of foreign keys.

(b) This part of the question seeks to determine if the candidate understands the concepts of entity integrity and referential integrity constraints and their importance. A brief description of each will suffice. If the description is backed up with an example, so much the better.

Five marks each for discussion of entity integrity and referential integrity constraints. For each, two marks for a definition, two marks for discussion of importance (entity integrity constraint needed to ensure addressibility of each tuple, referential integrity constraint for consistency of foreign key values), one mark for an example.

(c) This part seeks to determine if the candidate is aware of the implications of the two integrity constraints in terms of how the database is updated. A brief summary of the question of cascading or restricting changes to the database will suffice.

A description of the UPDATE/DELETE CASCADES/RESTRICTED constraint in SQL (five marks) with an example of its use (four marks).

Examiners' Guidance notes

a) Most people could provide an informal definition of a foreign key, and identify its role in database design. Most people provided an appropriate example. By and large this part was answered well.

b) Most people could describe refe rential integrity. Many people confused entity integrity with general application specific constraints on attribute values. Again, most people were able to state the importance of each.

c) This part of the question was well answered. Practically everyone identified the problems involved in maintaining these constraints in light of database updates.

Question 2

Answer Pointers

(a) This part of the question seeks to determine if the candidate is aware of the three schema (internal/conceptual/external) architecture of database management systems.

The answer should include a description of the internal, conceptual, and external schemas, the internal/conceptual and conceptual/external schema mappings (one mark each). An additional mark was given for an illustration of data independence.

(b) This part to the question seeks to determine if the candidate understands the concept of data independence, including the distinction between physical and logical data independence. The answer should include an example to demonstrate a practical understanding of the issues.

A definition of data independence in general was sought (two marks), followed by a definition of physical and logical data independence (two marks each). Three marks were allocated for an example illustrating each of (a) physical and (b) logical data independence.

(c) The final part to the question seeks to determine if the candidate understands how data independence is achieved and the relevance to the external/conceptual view mapping and the conceptual/internal view mapping.

This part of the question sought to determine if the candidate understood how physical and logical data independence are implemented by a database system. Three marks was allocated for a description of how physical data independence is provided by the relational data model and four marks were allocated for a description of how logical data independence can be provided by user views.

Examiners' Guidance notes

a) Most people described the three-level architecture well. No obvious problems arose with this part of the question.

b) Most answers demonstrated a good understanding of the concept of data independence. A number of answers confused physical and logical data independence. Some people did not know what data independence was at all.

c) Not very many people were able to answer this part of the question clearly. The answers were vague and confusing.

Question 3

Answer Pointers

a) i) The SQL code below in bold font implements the following steps:

Add a row for orderno=4 into the Orders Table:

Add a row for itemid=2 into the Items Table:

Add 2 rows (Items_Ordered) for orderno = 4 (one with itemid=2)
Add a row (Items_Ordered) for orderno = 5 (with itemid=2)

INSERT INTO Orders (orderno,custid,orderdate)
VALUES (4,99,'23-mar-00')
COMMIT

INSERT INTO Items (itemid,descr,qty_in_stock)
VALUES (2, 'spanner', 10)
COMMIT

INSERT INTO Items_Ordered (itemordered,orderno,qtyordered)
VALUES (1, 4, 1)
COMMIT

INSERT INTO Items_Ordered (itemordered,orderno,qtyordered)
VALUES (2, 4, 1)
COMMIT

INSERT INTO Items_Ordered (itemordered,orderno,qtyordered)
VALUES (1, 5, 1)
COMMIT

And similar for orderno=5, itemid = 2 etc

ii) The SQL UPDATE statement is required to change the contents of a table setting a column to a new value, based upon a condition:

UPDATE items SET unitprice = unitprice*1.175
WHERE itemid = 2
This updates price of itemid = 2 by 17.5%.

iii) The SQL delete statement is required using a subquery involving 2 tables to remove the selected item associated with orderno 4. Note that the Delete statement can only express joins by means of a Subquery ie:

DELETE FROM items WHERE itemid IN
(SELECT itemordered
FROM Items_ordered
WHERE orderno = 4
AND itemordered = 2)

Many candidates forgot to remove itemid=2 from the items table. An alternative interpretation was common in which candidates simply deleted the item from the items_ordered table, thus ignoring the consequences of maintaining the integrity of the Items Table. This interpretation made the query look like this:

DELETE * FROM itemsordered WHERE itemid = 2 AND orderno=4

This query did attract marks but many candidates forgot to update the items table to reflect the deletion of items from an orderline. It also needed:

update items set qty_in_stock = qty_in_stock . . . etc

b) A derivable column is total_qty_sold in the items table. It is normal not to store derivable data as the data will no longer be subject to dynamic updates unless a program or SQL code does this. When derivable data is stored it is usually because performance considerations outweigh the advantage of dynamically updateable data. Thus to calculate the value of the derived data identified above is:

```
SELECT SUM(qtyordered),itemorded
FROM itemsordered
GROUP BY itemordered
```

GROUP BY is used to form a set of values for each item.

Examiners' Guidance notes

Always use standard SQL (eg SQL92, SQL89) and avoid non-standard SQL dialects such as Microsoft-JET/ACCESS SQL.

The exact syntax of the INSERT statement was expected:

- valid SQL keywords (penalised for using ALTER,CREATE etc)
- the explicit reference to columns (minus 1 mark otherwise)
- prevent the INSERT of null values if stated Nulls='NO' in schema

Desirable features (omission was not penalised as strictly in this question):

- Correct reference to data types (ie quotes only for dates, text)
- Natural sequence of INSERTS applied (ie Order->Items_Ordered etc)
- Delimit statements with **COMMIT** (standard practice)
- Reasonable values for Qtyordered reflecting stock level

Question 4

Answer Pointers

a) The first part of the question seeks to determine if the candidate can produce an ER diagram.

Entities: Customer, Shipping Address, Product, Site, Order, Order Line (one mark each).

Attributes: (one mark for each entities attributes)

Customer(number, balance, credit limit, discount rate)
Shipping Address(address line, customer number(foreign key))
Product(number, description)
Site(product number (foreign key), quantity, danger level)
Order(order number, customer number (foreign key), address line, date)
Order Line(product number (foreign key), quantity, order number)

(b) The second part of the question seeks to determine if the candidate can transform an ER diagram into a relational database design.

A description of the RDB tables for the above diagram with appropriate primary and foreign key definition was sought here. (Four marks for table definitions and four marks for foreign key definitions.)

(c) The final part of the question seeks to determine if the candidate is aware of the pitfalls of blindly applying the theory of database design without looking at the practicalities.

The above design uses a separate table for the shipping addresses of the customers. This introduces extra complexity and overhead in looking up a customer's shipping address. If all customers had only one shipping address, this would not be necessary, the shipping address could be stored in the customer table. If only a few customers have more than one shipping address, it might be worthwhile storing a 'main' shipping address in the customer table and having additional shipping addressed in the Shipping Address table.

Three marks were allocated for noticing any problems in the design and two marks for suggesting a solution to such problems.

Examiners' Guidance notes

a) Generally this part of the question was well answered. Some of the diagrams were not very well drawn and were difficult to read.

b) Most people converted the ER diagram into tables well. Some candidates ignored the presence of multi-valued attributes, or created tables that would only record a single value for such attributes.

c) Only a few people were able to identify the drawbacks of their design. An understanding of the trade-offs involved in normalisation and database design was required to answer this part of the question.

Question 5

Answer Pointers

The report should be structured into 3 main areas. Indicative marks are shown:

1. Problem statement/assumptions, conclusions 6 marks
2. Technical Content (see note 1 below) 10 marks
3. Specific E-Commerce issues related to the scenario (see note 2 below) 9 marks

Note 1: Technical Content

- Suitability of DBMS and server support
- Software support from packages such as ASP/JDBC.
- Principle mechanisms CGI/Servlets
- Server side vs client side processing/scripting

- Connectivity via embedded SQL

Candidates were expected to have broad technical knowledge of the above features.

Note 2: Specific E-commerce problems raised in context to the scenario
At least one of the major problems should be covered, such as
- Security of connections,
- Confidentiality of users (encryption),
- Security and integrity of Corporate database – separate Database server?
- Support regarding backup/consistency

Examiners' Guidance notes
Only basic knowledge of Database-WWW interaction methods and techniques is expected at this level. However many candidates did not seem to be aware of the limitations of static HTML and the use of CGI or similar scripting techniques. The problem of state-less interactivity and static web pages apparent should be mentioned if CGI scripts are recommended. Other aspects missing include the dedication of the database server for web access. Very important issues including security mechanisms must be mentioned to give a balance to the obvious benefits of on-line transaction processing over WWW.

Question 6

Answer Pointers

a) Non-formal definitions acceptable, thus:

Determinant: an attribute (or group of attributes) that functionally determine one or more attributes in a relation. Attributes which are not determinants are normally functionally dependent on a particular determinant. (3 marks)
Candidate Identifier: an attribute (or group of attributes) that ensure each row in a relation is unique. Since there are usually many identifiers that could be chosen as an identifier the term 'candidate' is used. NB An Identifier can be implemented as a Primary Key. (4marks)

b) Examples:
Determinant: ResortName determines Area.
Candidate Identifier: ResortName + Hotel

(5 marks)

c) Non-trivial Determinants:

{Resort->Area}
{Hotel, Resort->Accom}
{Hotel, Resort->Price}
etc

Thus {Resort->Area}, {Hotel, Area->Resort} invalidate the BCNF rule.

Applying BCNF rule:

2 Tables for
Resort(PK)..Area
Hotel, Resort (PK).. Accom etc (7 marks)

d) Clearly the enterprise rule change effects the enterprise rules symbolically through FD's thus:

{Hotel, Resort->Price} to {Hotel, Resort, Nights, Month->Price}

Does this effect the Table structure - NO, but PK must be 4 attributes or defined as a new attribute (ie vacancy no).

(6 marks)

Examiners' Guidance notes

Candidates should be aware of and able to apply the two main different approaches to relational modelling:

- Formal step-wise approach UNF->1NF->2NF->3NF
- Informal approach based on BCNF and Functional dependency theory

This question uses the latter approach. This approach more fully tests understanding of the principles of relational modelling.

In summary this question revealed weaknesses in relational modelling. Many candidates understood the principles of normalisation but could not apply the intuitive approach required in this question despite the fact that candidates were essentially guided through the question. The crux of the question was working out the functional dependencies.