THE BRITISH COMPUTER SOCIETY

THE BCS PROFESSIONAL EXAMINATIONS
BCS Level 5 Diploma in IT

DATABASE SYSTEMS

18th October 2007, 10.00 a.m.-12.00 p.m.
Answer FOUR questions out of SIX. All questions carry equal marks
Time: TWO hours

*The marks given in brackets are **indicative** of the weight given to each part of the question.*

Calculators are **NOT** allowed in this examination.

1. Discuss and explain all the relevant concepts and constructs that must be addressed and implemented by a database programmer or database administrator when satisfying the following database issues:

a) Security constraints (on users) **(15 marks)**
b) Integrity constraints (on data) **(10 marks)**

You may use your own SQL code examples and any appropriate diagrams to support your answer to a) and b).

2. a) Using examples and diagrams where appropriate, define the following Relational Algebra Operators:
Project, Select, Union, Intersection, Join, Cartesian Product, Difference, Divide.
(16 marks)

b) Consider the following schema:

Employees (<u>Empno</u>, Ename, Dept, Job)
Skills (<u>Skill_no, Empno</u>)
Dept (<u>Deptno</u>, Deptname, No_Employees)

Write the relational algebra operators and steps required to compute the following sets:

i) List the names of the departments with more than 50 employees **(3 marks)**

ii) List the numbers and names of departments with at least one employee who has skill_no 5 **(6 marks)**

3. Suppose that a Housing Authority wants to use an Internet database application in its inventory of low-income housing units. Assume that the users of the application reside in a number of different government and non-profit organisations.

a) Sketch a possible three tiers system for supporting this Internet database application.
(10 marks)

b) Assuming that the data resides in an MS Access and that Windows 2000 is used on the Web Server, describe the functions of each of the tiers, you have produced in the previous question, and of the major components on each of the tiers. **(15 marks)**

4. Consider the following database scenario:

"A part manufacturing company has one engineering department in Manchester and two manufacturing plants, one in London and one in Hong Kong. Each part type is produced at only one manufacturing plant. Currently, the company has one database located in the Engineering department in Manchester. Applications at the manufacturing plants access this database via a communication network for whatever data they need.

One of the relations in this centralised database system is the PART relation, where data about the manufactured parts are kept. The attributes of the PART relation are: the part's number (Part#), part's name (Name), part's manufacturing cost (Cost), the part's drawing number that specifies its design (Drawing#), the name of the plant where the part is manufactured (Plant), and the quantity manufactured up to now (Qty).

An instance of the **PART** relation is the following:

PART					
Part#	**Name**	**Cost**	**Drawing#**	**Plant**	**Qty**
p2	Widget	200	123-7	London	500
p7	Gizmo	600	501-9	Hong Kong	1000
p3	Thing	100	238-2	Hong Kong	2000
p1	Gadget	1000	310-0	Hong Kong	40
p8	Acme	150	400-6	London	3000

The company has decided to move to a distributed database system where each of the sites has its own database."

i) Propose a fragmentation design of the PART relation that reflects the distribution of the company's sites and their functionality. **(8 marks)**

ii) Justify your proposed fragmentation design for the PART relation. **(5 marks)**

iii) Write an SQL statement for each of the fragments obtained in your design. **12 marks)**

5. a) Using specific examples from the Scenario (Appendix A) outline the main differences between a File-based approach to data processing with a Database-based approach to data processing. **(5 marks)**

b) Produce a logical database design for the scenario described in Appendix A. Your design MUST include each of the following:

i) An Entity-Relationship (ER) diagram using the Entities shown in bold font.

ii) A list of Table definitions that includes column names, primary keys and foreign keys.

You must state:

i) The ER diagram notation used, using examples from your ER model.

ii) The method that was used to transform your ER model into Tables.

iii) Any assumptions that you have made in your modelling.

(20 marks)

6. a) A travel agent has a late booking service where customers can book available seats on flights a day or so before the departure date. Sales are brisk as there are many opportunistic customers trying to get seats as discounted prices. Here’s a telephone conversation between the ticket agent and a customer.

Customer: "I'd like two tickets to Honolulu for January 15."
Ticket agent: "OK, I have two seats on flight number 192 leaving at 10 a.m."

Customer: "What is the price?"
Ticket agent: "£900"

Customer: "Great, I'll take them."
Ticket agent: "The screen has refreshed and the price now is £920"

Customer: "No problem please book them."
Ticket agent: "Oops. I just tried to reserve them, and they've been sold by another agent. Nothing else seems to be available on that date."

i) Explain the nature of the problem that has occurred above. **(5 marks)**

ii) Describe the mechanisms that a server-based DBMS would use to prevent the problem above. **(6 marks)**

b) Explain the function of the following program:

```
CREATE TRIGGER TestTrig2 ON TestTab FOR UPDATE AS
SAVE TRANSACTION MyName
INSERT INTO TestAudit
 SELECT * FROM inserted
IF (@@error <> 0)
BEGIN
  ROLLBACK TRANSACTION MyName
END
```

(8 marks)

c) Describe the differences between the RECOVERY mechanism in b) above and the RECOVERY mechanism that occurs when a DBMS crashes or is suddenly shut down. **(6 marks)**

Appendix A : Environment Agency Database for use in Q5

The Environment Agency (EA) is an organisation that is responsible for monitoring the level and flow rate of **Rivers**. They collect information so that they can predict flooding events and act on these appropriately.
The EA is organised into 5 **Regions** covering distinct geographical areas. Each region has a HQ that is responsible for the maintenance of rivers within its region.

The EA monitors about 500 rivers that have been identified as having a high risk of flooding. **Observatories** are positioned upstream on most of these rivers with about 5 observatories per river. The data collected from these observatories is sent to the EA Head Quarters where the data is processed and stored in an SQL database called the 'HQ database'.

Observatories contain equipment used to take **Measurements** of the following :-

- Precipitation (rainfall measured in cm)
- River levels (the maximum depth of water in a river measured in cm)
- Maximum Flow rate (measured in cubic cm per second)

The above measurements are recorded at discrete intervals on a 24x7 basis. When precipitation is above a certain threshold then measurements are taken approximately every minute otherwise they are taken approximately every hour.
Measurements for the last 20 years have been kept and these are stored in the HQ database.

There have been about 750 serious flooding events in the last 20 years and these are also stored in the HQ database. Each event lasts on average for 3 days before flood levels subside.

Another organisation called the Meteorological Agency (MA) send data transmissions known as '**Weather Feeds**' to the EA every 2 hours. A weather feed contains the following information:

- The date and time of the feed.
- The EA region that this data refers to
- Average recorded precipitation (in cm) since the last feed
- Predicted precipitation (in cm) for the next 24 hours

Information from Weather Feeds complements data from observatories and this is stored in the HQ database. Weather Feeds are particularly useful for rivers without an observatory or to cover for an observatory that is out of commission.

Figure A1 shows a snapshot of a large measurements table (called Measurements2005) which contains measurements recorded at EA observatories during 2005.

Measurements2005 contains the following columns :-
- Timestamp of data recording (when_taken).
- Precipitation (eg rainfall)
- River levels
- Observatory which recorded this data

The primary key of this table is the timestamp field ('when_taken')

FIGURE A1: A snapshot of a Table called Measurements2005

```
when_taken              precipitation river_level  ObservatoryID
---------------------------------- ----------- ------------
2005-07-17 12:33:43.003     1.0         23.3          123
2005-07-17 13:32:21.266     0.0         23.4          123
2005-07-17 12:34:51.772     1.1         43.4          295
2005-07-17 13:33:49.283     1.1         43.4          295
2005-07-17 14:31:34.033     1.7         24.5          295
```

2005-07-17 15:31:59.783	2.8	25.9	295
2005-07-17 15:38:52.843	1.1	43.4	123
2005-07-17 16:40:34.973	1.2	26.9	123
2005-07-17 17:31:22.983	0.2	27.2	295
2005-07-17 17:32:10.443	1.1	43.4	123
2005-07-17 18:33:04.203	0.6	27.2	295
2005-07-17 19:31:12.267	1.6	26.9	295
2005-07-17 20:32:59.553	1.8	27.5	123
2005-07-17 22:36:42.783	3.1	28.8	295
2005-07-17 23:38:49.208	2.6	29.9	295
2005-07-18 20:38:42.242	1.1	43.4	295
2005-07-17 21:36:21.281	3.8	28.3	123
2005-07-17 22:39:01.287	4.0	33.5	123
2005-07-18 00:00:00.296	4.9	38.8	123
2005-07-18 00:01:00.007	3.4	43.5	123
2005-07-18 00:03:03.124	2.6	44.0	123
2005-07-17 23:38:43.894	2.3	34.0	295
2005-07-18 00:04:45.005	1.0	44.1	123
2005-07-18 00:05:49.255	1.2	44.5	123
2005-07-18 00:06:51.562	0.9	42.5	123

Turn over]

**** END OF PAPER ****