

THIS PAPER IS NOT TO BE REMOVED FROM THE EXAMINATION HALLS
--

UNIVERSITY OF LONDON

291 0209 ZA

BSc Examination
for External Students

**COMPUTING AND INFORMATION SYSTEMS AND
CREATIVE COMPUTING**

Database Systems

Dateline: Tuesday 12 May 2009 : 2.30 – 5.30 pm

Duration: 3 hours

This paper consists of five questions. Each question carries **25** marks. Candidates should answer **FOUR** questions. Candidates may choose **any four** questions. Full marks will be awarded for complete answers to **FOUR** questions.

The mark carried by each part is printed within square brackets. **Gauge the time to be spent on each part by the number of marks awarded.**

No calculators may be used.

Question 1

A. What do we mean by the terms (1) "Data Manipulation Language" and (2) "Data Definition Language"? For each, give two different examples of SQL statements which illustrate them. [6 marks]

B. Most database management systems have one or more files which make up a "data dictionary" or "system catalogue". Describe the sorts of data which might be found in a system catalogue, and briefly indicate to what uses each sort of data might be put by the database system. [6 marks]

C. Database theorists often mention the external schema, the conceptual schema and the internal schema when referring to a database. What is meant by these three terms? [6 marks]

D. (1) In the context of a distributed database system, what is (a) horizontal and (b) vertical fragmentation? [3 marks]

(2) A bank keeps data on its clients in a table, where the columns have the following meanings: *ClientNum* is a unique identifier for each client, and *ClientName* is the client's company name. *Branch* is the bank branch that deals directly with a client; *AcctBalance* and *CreditRating* describe the financial status of a client. *Contact* and *Phone* refer to someone who works for the client who is the branch's contact. Some example tuples are shown below.

CLIENTS

Primary Key: Clientnum

<u>Client Num</u>	Client Name	Branch	Acct Balance	Credit Rating	Contact	Phone
0987224	Smith Bros	Wentworth	34556.50	5	Bill Smith	44566
7686334	Axion	Arcade	933.75	3	Oksana Poltava	76443
6452467	Fleuria	Wentworth	2309.00	5	Hua Hsu	33216
9373544	ProPon	Wentworth	3445.00	5	Jessica Jones	32468
3759334	Sandino	Arcade	2230.50	4	Heidi Schonberg	35569
8704373	Xon Corp.	Arcade	56.00	2	Kim Jong Pak	33324

The company has a central HQ, which needs to access client *AccountBalances* and *CreditRatings*. Its branches seldom need to access those fields, but often need to access contact names and phone numbers of the clients who are their responsibility. Describe how this table could be horizontally and vertically fragmented in a distributed database system, where every branch, and the central HQ, has a computer, all linked by a network.

[4 marks]

[Total 25 Marks]

Question 2

A. Concurrent access to the same records by two or more different processes, where one or more of those processes can update the records, can give rise to problems. Describe one of these possible problems and illustrate how the problem can be prevented by using a record locking mechanism.

[10 marks]

B. What is meant by "query optimisation"? Give an example of a query involving both a join between two different relations and selection statements of the relations, and show how this query could be executed in two different ways, one of them less efficient than the other. What determines the degree to which the more efficient algorithm is better than the less efficient one?

[5 marks]

C. A new business wants to create a database which will hold information on its employees. Employees are identified by unique Employee Numbers, which are always exactly twelve characters long. They have Surnames, Middle Names, and First Names, each of which can be up to 24 characters long. Each employee must have a Surname and a First Name, but not necessarily a Middle Name. Employees may (or may not) also have Job Titles, which take up to 36 characters; and Salaries, which must be more than 20,000 a year.

The business also wants to record information on the children of each employee, if they have them. It wants to record the child's First Name, which will take up to twelve characters; its sex, which must be either 'M' or 'F', and its date of birth. All children in the database must have a parent who is an employee. If an employee quits, the company does not keep the records of his children. (In other words, the child is a 'weak entity type,' dependent on a parent employee.)

Write the SQL statements that will create the tables which can hold this information. Only two CREATE statements will be needed. Be sure to include any constraints listed above.

[10 marks]

[Total 25 marks]

Question 3

A newly-formed ferry company will run ferries from several ports in England, to several ports on the European continent. Its database needs to record information about its Ferries, Connections, Voyages and Customers.

Connections are regularly-scheduled sailings. They are identified by Connection Numbers. Each Connection has a Departure Port and a Destination Port, a Departure Time and an Arrival Time, and a Person-Ticket-Price and a Personal-Ticket-with-Automobile-Berth Price. (Some people travel by foot, others bring their cars.)

A Customer is given a unique Customer-Number, has a Family-Name and First-Name, and may have a Passport Number. He can book one or more connections, each for a specific date. A given connection on a particular date is called a voyage.

Each voyage will have many customers, and one Captain and one particular ferry. A customer's booking of a voyage will be for either a personal seat alone, or a personal seat plus an automobile berth. A given connection is in one direction between two cities. (Return journeys have different connection numbers.)

For example, Connection DC05 takes place between Dover and Calais, sails at 9.30 am on Monday through Saturday, and is always made by a SeaMaster-B type ferry. (The ferry company owns several SeaMaster-B ferries, and any one of them may be used on this Connection on a particular date. A captain may switch between ferries, commanding one vessel on Monday and a different one on Tuesday.)

Ferries will be a certain Model (such as a SeaMaster-B), be identified by a unique Vessel-Identifier, and will have a certain number of seats and a certain number of Automobile berths. (These may differ slightly from Ferry to Ferry, even among Ferries of the same model.)

An example of a typical transaction might be that Customer P2094 booked a place for one person on Connection DC05 for the 17th of March 2008. On this Voyage he was allocated Seat 5 in Row K. The Ferry used for that particular voyage had the Vessel-Identifier UKF6622, and was a SeaMaster-B, commanded by Captain Pierre Lamartine, whose Employee Number is U998A. This particular vessel had room for 75 automobiles and 450 passengers.

A. Draw an Entity/Relationship diagram to represent this situation. Include only the entity types and relationships, but not the attributes.

[12 marks]

B. Design a relational schema which could hold the information represented by the E/R diagram you drew up for **Part A**, and the attributes mentioned in the description of the question.

[13 marks]

[Total 25 marks]

Question 4

A database was created using the following statements.

```
CREATE TABLE Specialist (  
    SpecialistID      CHAR(8)  
    Surname           VARCHAR(32),  
    FirstName         VARCHAR(32),  
    Division          VARCHAR(12),  
    SpecialismCode    CHAR(8),  
    PRIMARY KEY       (SpecialistID),  
    FOREIGN KEY       (SpecialismCode) REFERENCES Specialism(SpecialismCode)) ;
```

```
CREATE TABLE Specialism (  
    SpecialismCode    CHAR(4),  
    Description        VARCHAR(80),  
    HourlyCharge      INT,           -- in Dollars  
    PRIMARY KEY       (SpecialismCode) ;
```

```
CREATE TABLE ProjSite (  
    ProjSiteNum       INT,  
    DateBegun         DATE,  
    DateEnded         DATE,  
    Budget            INT,  
    Overseer          CHAR(8)  
    PRIMARY KEY       (ProjSiteNum),  
    FOREIGN KEY       (Overseer) REFERENCES Specialist(SpecialistID) );
```

```
CREATE TABLE Assignment (  
    SpecialistID      CHAR(8)  
    ProjSiteNum       INT,  
    AssignmentDate    DATE,  
    HoursWorked       INT,  
    PRIMARY KEY       (SpecialistID, ProjSiteNum, AssignmentDate),  
    FOREIGN KEY       (SpecialistID) REFERENCES Specialist(SpecialistID),  
    FOREIGN KEY       (ProjSiteNum) REFERENCES ProjSite(ProjSiteNum));
```

SPECIALIST

SpecialistID	Surname	Division	SpecialismCode
--------------	---------	----------	----------------

Specialism

SpecialismCode	Description	HourlyCharge
----------------	-------------	--------------

ProjSite

ProjSiteNum	DateBegun	DateEnded	Budget	Overseer
-------------	-----------	-----------	--------	----------

Assignment

SpecialistID	ProjSiteNum	AssignmentDate	HoursWorked
--------------	-------------	----------------	-------------

(question continues on next page)

A Express the following natural language queries in SQL.

(1) List the codes and descriptions of all specialisms with an hourly rate greater than 175 Dollars per hour, in alphabetical order of description.
[1 mark]

(2) List the SpecialistID and Surname of all Specialists who have the Specialism described as 'Acoustic Surveying' who work in the 'Special Projects' Division.
[2 marks]

(3) How many Specialists have the Specialism 'Acoustic Surveying'?
[2 marks]

(4) List all ProjSite numbers and the total numbers of hours worked by Specialists at that ProjSite, for ProjSites where the total time worked by Specialists has been more than 1000 hours.
[2 marks]

(5) For all the ProjSites that were active on '1 January 2005' (i.e. the start date was before this date and the end date was after this date) list the Specialist Surnames, ProjSite numbers, date and number of HoursWorked, ordered by ProjSite number and within that by Specialist Surname and within that Surname by date, for all Specialists who worked in the ProjSites.
[3 marks]

(6) List all Specialists (Surname and Division) with an hourly charging rate greater than the average hourly charging rate.
[3 marks]

B. What SQL statements would you need to add the following constraints to the tables created in the previous part of this question?

(1) The maximum number of HoursWorked that a Specialist can be allocated to a ProjSite, per day, is 8; the number of HoursWorked has to be positive.
[2 marks]

(2) "Surname" is a candidate key in Specialist. [1 mark]

(3) **(a)** If the SpecialistID is changed for a Specialist who is the Overseer of a ProjSite, this update must be propagated in the table ProjSite
(b) a Specialist cannot resign (delete the tuple from Specialist) if he is managing a ProjSite (note that a ProjSite will be deleted from the database if it is closed down).
[3 marks]

(4) The hourly rate for 'Acoustic Surveying' should be greater than the average rate.
[3 marks]

(5) The Overseer of every ProjSite must be from the 'Management' Division.
[3 marks]

[Total 25 marks]

Question 5

A. The following table records the results for would-be actors who are undergoing a preliminary screening audition for a particular role in a play. If an actor is turned down for a role, in the preliminary screening, he is not allowed to re-audition for that role again. If he passes, he will be re-auditioned later along with others who passed the preliminary screening.

Actor	Role	Audition Date & Time	Result
John Jones	Willy Loman	21 July 0900	reject
John Jones	Ben Loman	21 July 1130	pass
Gerry Olavsky	Willy Loman	21 July 0900	pass
Milton Ntobi	Willy Loman	21 July 1000	pass

The primary key of this table is Actor + Role .

(1) What bad consequences could follow if we mis-identified the primary key of the table as

(a) Actor alone?

(b) Actor + Role + Audition Date & Time?

[4 marks]

(2) Suppose it is decided to allow actors to re-audition for roles for which they were initially rejected, at a later date and time. Would we need to change the definition of the primary key? Explain your answer.

[1 mark]

B. Consider the following table, which records the email addresses and phone numbers of employees. (Employees are identified by Employee numbers) Note that the employee whose employee number is E933 has one email address and two phone numbers, and the employee whose employee number is E204 has two email addresses and one phone number.

EMPLOYEE DETAILS

Employee	Email address	Phone number
E933	John22@Nomail.com	309-8219
E933	.null.	7648099432
E204	SueTwo@netmail.com	309-3304
E204	Susie222@Mymail.com	.null.

Although this relation does not violate the rule "let every determinant be a candidate key", yet it is a very poor design. Explain why, and suggest an alternative way to represent the information held by it.

[5 marks]

The following table applies to Parts C, D and E. It records information about the annually-recorded weight of particular prize sheep. Each sheep belongs to a particular owner. No sheep is owned by more than one owner. Each sheep's birthdate is recorded. Every year, a veterinarian weighs each prize sheep and records its weight. The date of the weighing, and the ID number of the vet is recorded, plus the pager number of the vet.

(question continues on next page)

SheepID	Owner	Birthdate	Weighing Date	Vet	Weight	VetPager Num
K3922	McNab013	23 May 96	12 Dec 1998	M330	52	7633088852
K3922	McNab013	23 May 96	23 Oct 1999	S929	54	7609865463
K3922	McNab013	23 May 96	7 Nov 2000	M330	53	7633088852
K3922	McNab013	23 May 96	8 Jan 2001	P301	53	7682907965
K3922	McNab013	23 May 96	29 Nov 2002	P301	52	7682907965
K3922	McNab013	23 May 96	15 Nov 2003	S929	51	7609865463
T8832	McNab013	3 May 95	12 Dec 1996	K339	49	7682907965
T8832	McNab013	3 May 95	10 Dec 1997	S929	50	7609865463
T8832	McNab013	3 May 95	23 Oct 1998	K339	50	7682907965
T8832	McNab013	3 May 95	30 Dec 1999	K339	51	7682907965
T8832	McNab013	3 May 95	3 Nov 2000	T975	52	7633088852
P9742	Smith002	29 Apr 95	22 Sep 1996	K339	54	7682907965
P9742	Smith002	29 Apr 95	19 Oct 1997	S300	55	7629920821
P9742	Smith002	29 Apr 95	24 Sep 1998	K339	55	7682907965
P9742	Smith002	29 Apr 95	8 Nov 1999	S929	54	7609865463
P9742	Smith002	29 Apr 95	24 Nov 2000	S929	53	7609865463
M3110	Smith002	23 May 96	22 Sep 1997	S300	57	7629920821
M3110	Smith002	23 May 96	30 Sep 1998	S300	58	7629920821
M3110	Smith002	23 May 96	18 Oct 1999	K339	58	7682907965

C. Identify the Functional Dependencies in this table. [3 marks]

D. This table is susceptible to update, deletion, and insertion anomalies. Give an example, based on the table, of each kind.

[6 marks]

E. Bring the table to BCNF, specifying the Primary Keys of each table, and showing the extension of the resulting relations.

[6 marks]

[Total 25 marks]

END OF EXAMINATION