UNIVERSITY OF LONDON

291 0209W

BSc Examination 2006

For External Students

COMPUTING AND INFORMATION SYSTEMS

CIS209 Database Systems (Western)

Duration: 3 hours

Date and time: Tuesday 9 May 2006: 10.00 - 1.00pm

This paper consists of **five** questions. Each question carries 25 marks. Answer **four** questions only. Full marks will be awarded for **complete** answers to **four** questions.

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

THIS EXAMINATION PAPER MUST NOT BE REMOVED FROM THE EXAMINATION ROOM.

QUESTION 1

a) Create a conceptual data model – ER diagram – for the Fit-Gym Fitness club describe below. [8]

b) If your solution contains multi-valued or composite attributes, many-to-many relationships or relationships with attributes, transform it into an equivalent one that does not contain these aspects. State any further assumptions you make, different from the ones explicitly made in the text (if you make any). If your solution does not contain such aspects, move to point c). [9]

c) Suppose that each apparatus has associated a set of guidelines that provides for a given sex and age the number recommended weight and number of repetitions. Show how the above ER diagram should be amended in order to accommodate this. [8]

The manager of the Fit-Gym fitness club wants to develop a database system which to facilitate the booking and renting of the equipment available in the gym. You are asked to create an ER diagram for the following specification.

The data stored in the database should represent information about equipment and the members of the club. For each member the database should store: the membership number, name, age, sex (male or female), address (number, street, postcode), telephone number and date joined. For each type of equipment/apparatus the database should store the: name, type, the number of apparatuses of that name and type available in the club, the cost per hour to use, the cost per day to rent (if available), the group of muscles it is useful for and the things it is counter-indicated for. Both last attributes can have more than one value. For instance, a certain type of apparatus works out both biceps and triceps and can be counter-indicated for upper-back problems and high blood pressure.
Member can book different apparatuses in advance; for this they will have to specify the day, starting time and ending time. Certain types of equipment can also be rented out. When such a transaction is made, the date of the transaction and the date when it is supposed to be returned should be recorded. Only one type of equipment can be rented at any one time by a member.

QUESTION 2

Consider the relation Law-Advice presented below, used by a law firm. The relation is already in first normal form (1NF). It contains information about
- the clients of the firm: "Customer" and "Address" represent the client's name and the address respectively; "Problem" and "Reason" represent the problem for which they needed legal advice and the reason they provided for the problem;
- "Lawyer" and "Office" represent the name of the lawyer whom the custom has been assigned to and the office where the lawyer works, respectively;
- if the problem requires, then a hearing will be assigned for the client; the hearing is identified by a "Hearing-ID" and takes place at a certain date and venue, represented by "Date" and "Venue" respectively.

Consider the following assumptions
- there are no two clients with the same name;
- only one reason can be provided for a given problem by a certain person;
- a lawyer uses more than one office, but an office can only use by one lawyer;
- a client, for a certain problem, is referred to one office only;
- a client, for a certain problem, can go to just one single hearing (that is if the hearing is required);
- more than one client can be sent to the same hearing.

a) This table is susceptible to update anomalies. Provide examples of how insertion, deletion and modification (update) anomalies can occur in this table. [6]

b) Identify the functional dependencies existing in this table, based on the assumptions presented above. [5]

c) Bring the table to BCNF (you can do this directly; you do not have to go through intermediate forms (i.e. 2NF and 3NF); however, if you find it easier, you can go through intermediate forms). Specify the primary keys and the alternate keys (if any) for all the resulting relations. Show the extension of the resulting relations as well. [14]

Client	Address	Problem	Reason	Lawyer	Office	Trail-ID	Date	Venue
P. Linch	SE14 6NW	tax	forgotten to declare income	P. Ray	RoomA, Hampton Place	TA1552	23/05/00	Court1, NE14 6WW
P. Linch	SE14 6NW	domestic	quarrel with neighbour	P. Wolf	Room01, Sloan Square	–	–	–
M. Ross	NE25 6UU	violence	lost temper	P. Wolf	Room02, Sloan Square	VI09987	04/07/00	Main Court, Sloan Square
M. Ross	NE25 6UU	abuse	lost temper	P. Wolf	Room03, Sloan Square	VI07889	04/07/00	Main Court, Sloan Square
M. Ross	NE25 6UU	tax	did not pay	P. Ray	Room10, Sloan Square	TA8990	12/09/00	Court1, NE14 6WW
R. Wild	CO98 7PQ	tax	forgotten to declare income	M. Philips	RoomB, Hampton Place	TA1552	23/05/00	Court1, NE14 6WW

QUESTION 3

Consider a company that holds information about its employees, departments, projects and number of hours spent by each employee on each project. Each department has one single manager. Each project belongs to a single department. One department can have more than one project. Each project has a project manager, who does not have to be the manager of the department.
The following definitions exist in the database (the name of the relations and attributes are self explanatory):

```
CREATE TABLE Employee (
     Emp_id              INT,
     Name                VARCHAR(30),
     Address             VARCHAR(100),
     Age                 INT,
     Dept_id             VARCHAR(20),
     PRIMARY KEY         (Emp_id),
     FOREIGN KEY         (Dept_id) REFERENCES Department);
CREATE TABLE Department (
     Dept_id             VARCHAR(20),
     Name                VARCHAR(50),
     Manager             INT,
     Budget              INT,
     PRIMARY KEY         (Skill_code),
     FOREIGN KEY         (Manager) REFERENCES Employee (Emp_id));
CREATE TABLE Project (
     Proj_no             INT,
     Proj_name           VARCHAR(100)
     Proj_manager        INT,
     Dept_id             VARCHAR(20),
     PRIMARY KEY         (Proj_no),
     FOREIGN KEY         (Proj_manager) REFERENCES Employee(Emp_id),
     FOREIGN KEY         (Dept_id) REFERENCES Department);
CREATE TABLE Worked_on (
     Emp_id              INT,
     Proj_no             INT,
     Hours_worked        INT,
     PRIMARY KEY         (Emp_id, Proj_no),
     FOREIGN KEY         (Emp_id) REFERENCES Employee,
     FOREIGN KEY         (Proj_no) REFERENCES Project) ;
```

a) Express the following natural language queries in SQL.

1) List all employees (name and address) who are younger than 30, in alphabetical order. [2]

2) List all employees (all details) who work for the 'Development' or 'Support' department. [3]

3) List how many staff work for each department that has the budget below 1000,000 (number and name of department); order the list alphabetically on department name. [3]

4) Produce a list of all the projects and their managers (name) for the projects that have as manager a departmental manager (not necessarily the one that the project belongs to) and order the list on project names. [3]

(Question continues on next page)

5) List the projects and the total number of hours worked for, for all the projects that were supervised by the manager having the id "MA105". [3]

6) List the departments (name) and their manager (name), for all the departments whose budget is greater than the average departmental budget (the average departmental budget is represented by the sum of all departmental budgets divided by the number of departments). [3]

b) Express the following constraints in SQL.

1) The maximum departmental budget is 5000,000; the budget also has to be positive. [2]

2) "Proj_name" is a candidate key in Project. [2]

3) If the Dept_id is changed for a department that "owns" some projects, this update must be propagated in the table Project (a department is said to own a project if their are projects that belong to that department); if a department is closed (i.e. deleted from the Department table) then all the projects it owns must also be deleted from Project. [4]

QUESTION 4

a) State the data access protocol in the context of concurrency control and illustrate how it can be applied in order to solve a concurrency problem (the choice of the concurrency problem –lost update, uncommitted dependency or inconsistent analysis – is yours). Can there any other problems arise as a result of applying the data access protocol? Is there a solution to them? [14]

b) State the main objective (or fundamental principle) of distributed database systems. Then, explain what is it meant by replication independence (you can provide an example), describe the advantages provided by data replication and state the replication independence principle. [11]

QUESTION 5

a) Consider the following base relations (only the heading of the table is provided below):

Student

Username	Name	Address	Photo	Personal_tutor

Primary key: Username (the name used to log into the database system)
Alternate key: Name
Foreign key: Personal_tutor (references the table of tutors)

Exam_Result

Username	Exam	Result

(Question continues on next page)

Primary key: (Username, Result)
Foreign key: Username (references Student) and Exam (references the table of modules)

Write the SQL statements for the following security rules (hint: for some rules, the use of views is indicated):
1) Allow everybody to see the name and the photo of each student. [3]
2) Allow John Creek (username: mcsjk) to modify his own name, address and photo. [3]
3) Allow the tutor Peter Clark (username mcsspc) to see the name, address and photograph of all his personal tutees. [3]
4) Allow John Creek (username: mcsjk) to see his exam results. [3]
5) Allow students to see their own results only (hint: use the function "user()" that returns the username of a current user). [3]

b) Discuss the idea of query processing in distributed databases using the example given below. [10]

```
Consider three relations
Drugs(Drug-ID, D-Name, Manufacturer, Recommended-For)
Pharmacies(Pharmacy-ID, P-Name, Location)
Availability(Drug-ID, Pharmacy-ID, Stock) - existing stock of drugs
in each pharmacy
Suppose these relations are implemented in a distributed database
system. Suppose that the first relation - Drugs - is stored at site A
and the last two relations - Pharmacies and Availability - are stored
at site B. Suppose the following query is issued at site C (select
all the pharmacies in York that have in stock drugs recommended for
ulcer)
SELECT    P-Name, Location
FROM                              Drugs D, Pharmacies P, Availability A
WHERE     D.Drug-ID = A.Drug-ID AND P.Pharmacy-ID = A.Pharmacy-ID AND
                      Recommended-For = 'ulcer' AND Location = 'York';
Assume that
- each tuple is 2000 bits long;
- Drugs, Pharmacies and Availability have 400000, 800 and 8000000
tuples, respectively.
- estimated number of pharmacies in York is 20
- data transfer rate is 100,000 bits per second.
Describe at least two possible algorithms of processing this query
and illustrate the difference in processing time
(You will need to make some other assumptions apart from the ones
made above).
```

END OF EXAMINATION