UNIVERSITY OF LONDON

291 0209E

BSc Examination 2006

For External Students

COMPUTING AND INFORMATION SYSTEMS

CIS209 DATABASE SYSTEMS (EASTERN)

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 Video Club of Silversmiths College described 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) Transform the ER diagram into a relational model, specifying the primary, alternate (if any) and foreign keys. [8]

The Information Services of Silversmiths College have decided to create a Video Club for both students and staff. The members of this club will be allowed to either borrow videos for home or watch them using a set of television sets installed in the building. The present database that is in use for the library (for borrowing books) will have to be extended to cover this new video facility. For this, you are required to create an ER diagram. Work as if the old system did not exist, i.e. consider only the present application.

The data stored in the database should represent information about videos and members. The data required for videos includes: title, category (comedy, drama, horror, classic, musical, science fiction, science, ...) director, main actors, status (used only when the video is issued to a member) cost and the total number of copies that the library holds. A video can have more than one category (e.g. classic and drama) and also more than one main actor. A member of the video club is described by the following attributes: membership number, name (first and last), date joined, address and telephone number. Both staff and students can be members of the Video Club.
Videos are of two categories: for hiring (i.e. they can be taken at home for a period of maximum seven days) or for borrowing (i.e. they can be taken inside the library only and used with the television sets provided). Each member may hire an unlimited number of videos, as long as they are available. For each video hired, the date-out and the date-in must be recorded. The date-in depends on the status of the video and whether the member is a student or staff; it is not computed automatically, but it is entered by the issue desk. Each member may borrow at most one video. For each video borrowed, only the time of issue must be recorded.

QUESTION 2

Consider the relation Patient-Treatments presented below. This relation is already in first normal form. "Age", "Address" and "Disease" represent the age, address and disease of a patient. "Speciality" represents the doctor's speciality. "Treatment" is the code of the treatment administered to a certain patient for a certain disease and is described by a set of drugs – "Drug"– taken at certain specified times – "Time". Consider the following assumptions:

- no two patients have the same name (represented by the attribute "patient");
- a patient can suffer from more than one disease;
- for each disease, a patient is assigned a unique treatment room;
- a treatment room is only used by one doctor;
- a doctor can use more than one treatment room;
- a patient is given a single treatment for a certain disease;
- a disease does not have a single treatment associated with it;
- a treatment has one single main drug;
- two or more treatments can be based on the same main drug;
- each drug has a unique administration mode (in the form of a text, therefore atomic).

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]

Patient-Treatment

Patient	Disease	Age	Address	Doctor	Treatment Room	Treatment	Main-Drug	Administration
M. Jackson	ulcer	56	London, SE14	M. Stevens	Room10	UL100	a220	2x after meals
M. Jackson	high blood pressure	56	London, SE14	P. Wolf	Room01	PP100	Betamicin	once a day
M. Jackson	high cholesterol	56	London, SE14	M. Brick	Ward A	Diet	–	–
J. Peters	stomach ache	30	London, SW12	M. Stevens	Room12	UL100	a220	2x after meals
J. Peters	head ache	30	London, SW12	M. Brick	Ward C	PP100	Betamicin	once a day
R. Philip	ulcer	40	London NE21	P. Lomu	Room16	NN25	Amophilin	1x after meals

QUESTION 3

Consider a software engineering company that takes on projects and uses its staff to carry out these projects. Each member of staff has one strong skill. The company has a payment scheme for the set of skills its staff have. A project has just one project manager. Members of staff are allocated periods of time to work on the project. The following definitions exist in the database (the name of the relations and attributes are self explanatory):

```
CREATE TABLE Staff (
     Staff_no            INT,
     Name                VARCHAR(30),
     Department          VARCHAR(20),
     Skill_code          CHAR(5),
     PRIMARY KEY         (Staff_no),
     FOREIGN KEY         (Skill_code) REFERENCES Skill_payment ) ;
CREATE TABLE Skill_payment (
     Skill_code          CHAR(5),
     Description         VARCHAR(80),
     Charging_rate       INT,        -- Pounds per hour
     PRIMARY KEY         (Skill_code) ) ;
CREATE TABLE Project (
     Project_no          INT,
     Start_date          DATE,
     End_date            DATE,
     Budget              INT,
     Manager             INT,
     PRIMARY KEY         (Project_no),
     FOREIGN KEY         (Manager) REFERENCES Staff(Staff_no) );
CREATE TABLE Allocation (
     Staff_no            INT,
     Project_no          INT,
     Date_worked_on      DATE,
     Hours               INT,
     PRIMARY KEY         (Staff_no, Project_no, Date_worked_on),
     FOREIGN KEY         (Staff_no) REFERENCES Staff,
     FOREIGN KEY         (Project_no) REFERENCES Project) ;
```

a) Express the following natural language queries in SQL.

1) List all skills with a charging rate greater than 60 Pounds per hour, in alphabetical order of description. [2]

2) List all staff with the skill description 'Programmer' who work in the 'Special Projects' department. [3]

3) How many staff have the skill 'Programmer'?. [3]

4) For all the projects that were active on '1 January 2000' (i.e. the start date was before this date and the end date was after this date) list the staff name, project number, date and number of hours, ordered by project number and within that by staff name and within that name by date, for all the staff who worked on the active projects. [3]

5) List all project numbers and the total numbers of hours worked on, for projects that were worked on for more than 100 hours? [3]

6) List all staff (name and department) with a charging rate greater than the average charging rate (the average charging rate is the sum of all charging rates for all different skills divided by the number of skills). [3]

b) Express the following constraints in SQL.

1) The maximum number of hours that a staff member can be allocated to a project, per day, is 10; the number of hours has to be positive. [2]

2) "Name" is a candidate key in Staff. [2]

3) If the Staff_id is changed for a staff member that is the manager of a project, this update must be propagated in the table Project; a staff member cannot resign (delete the tuple from Staff) if s/he is managing a project (clarification: a project is deleted from the database as soon as it finishes). [4]

QUESTION 4

a) Security restrictions can have different origins: legal, ethical, political, strategic, etc. Discuss three of them (not necessarily on the list above), considering concrete example (such as the database of a hospital). In one case, give an example of a security rule, either is pseudo-code or in SQL, to illustrate your ideas. [11]

b) Discuss the concept of data fragmentation and explain the fragmentation independence principle (in the context of distributed database systems). [14]

QUESTION 5

a) Describe one of the possible problems generated by concurrent access to the same data (the lost update problem, the uncommitted dependency problem, the inconsistent analysis problem) and illustrate how the problem can be resolved if the locking mechanism is utilised. Can there additional problems occur? How could they be resolved? [11]

b) Discuss the idea of query processing in distributed databases using the example given below. [14]

```
Consider three relations
Books(Book-ID, Title, Author, Year)
Libraries(Lib-ID, Location)
Availability(Book-ID, Lib-ID) - existing books in libraries
Suppose these relations are implemented in a distributed database
system. Suppose that the first two relations - Books and Libraries-
are stored at site A and the last relation - Availability - is stored
at site B. Suppose the following query is issued at site C (select
all the books published in 1999 stored at the libraries in London)
SELECT    *
FROM      Books B, Libraries L, Availability A
WHERE     B.Book-ID = A.Book-ID AND L.Lib-ID = A.Lib-ID AND
                              Year = '1999' AND Location = 'London' ;
Assume that:
- each tuple is 1000 bits long;
- Books, Libraries and Availability have 100000, 100 and 5000000
tuples, respectively.
- estimated number of books in 1999 is 5000
- 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 may need to
make some other assumptions apart from the ones made above).
```

END OF EXAMINATION