THIS PAPER IS NOT TO BE REMOVED FROM THE EXAMINATION HALLS

UNIVERSITY OF LONDON **291 0209 ZB**

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 College of Agricultural Sciences offers various courses leading to a BSc.

Each course – which has a unique name -- has its own set of criteria for admission (a special entrance exam, or passing the International Baccalaureate , or science and maths A-Levels) – all these represented as a single text entry in the database; its own duration (some take three years, some four, and some five); and its own method of allowing students to gain practical experience (some via work experience, others by a practical session during the summer at the College's farm). A course can exist in the database without any students being enrolled in it.

Every course consists of a set of modules, some of which are compulsory, and some of which are optional , each carrying a certain number of course credits. Each module has a name, a syllabus, and the modules, if any, which are pre-requisite for it. A particular module is taught by one lecturer, and in a given year, a particular lecturer might teach none, one, or several modules. Note that a given module can be compulsory for one or more courses, and optional for one or more courses.
We also want to record which lecturers can teach which modules – a lecturer may be recorded as being able to teach a module which he is not, however, teaching that year.
A student in the database must be registered for one, and only one, course. Optional modules, during a given year, may have no students registered for them. Courses may share modules – that is, two or more courses may require students to take the same module.

Students are identified by a Student Number. We want to record the course they are enrolled for, the year they enrolled, the modules they have already attempted, the year they were examined in that module, and the grade they received. (Students may re-sit examinations in modules they fail.) A student can be registered for a course, but not yet registered for any optional modules.

Lecturers are identified by Lecturer ID's. We want to record their names, birthdates, the modules they can teach, and the modules they are teaching currently. It is possible to have lecturers who are not teaching any modules currently, but all lecturers must be able to teach at least one module.

Modules are identified by a Module Name, and we want to record the number of credits a module is worth. A new module which has not yet started may not have any lecturers able to teach it, but will still be represented in the database. No module will be taught by more than one lecturer.

A. Draw an Entity/Relationship Diagram that expresses the relationships among the entity types described above. You need not indicate the

(question continues on next page)

attributes of each entity type, but you should show "participation and cardinality constraints". (In other words, show whether a relationship is necessary or optional for each participating entity type and if there are minimum or maximum numbers of participating entities in a relationship.)

[11 marks]

B. Design a fully normalized relational schema that can capture the data relationships expressed in your Entity-Relationship diagram. Be sure to indicate the primary and foreign keys of each relation.

[14 marks]

[Total 25 marks]

Question 2

The following table records information about the percentage of ground cover found in certain uniquely numbered plots of land, which are being observed over time to track the growth of vegetation on them. Each plot is found in a given district, and no plot overlaps districts. Each plot's size in hectares is recorded. Roughly every two years, an on-site inspection is done to record the amount of ground cover growing on the site. The date of the inspection, and the identification number of the inspector, is recorded, along with the inspector's telephone number.

PlotNo	District	Hectares	Inspection Date	Inspector	Ground Cover (%)	Phone#
2455	West Lake	32.5	12 Dec 1998	R409	34	8097733027
2455	West Lake	32.5	23 Oct 2000	P528	32	7609865463
2455	West Lake	32.5	7 Nov 2002	M330	32	7633088852
2455	West Lake	32.5	8 Jan 2005	P301	35	7682907965
2455	West Lake	32.5	29 Nov 2006	P301	36	7682907965
2455	West Lake	32.5	15 Nov 2008	P528	36	7609865463
3891	West Lake	12.3	12 Dec 1998	P301	38	7682907965
3891	West Lake	12.3	10 Dec 2000	P528	36	7609865463
3891	West Lake	12.3	23 Oct 2002	P301	36	7682907965
3891	West Lake	12.3	30 Dec 2004	P301	38	7682907965
3891	West Lake	12.3	3 Nov 2006	M330	37	7633088852
2349	Bluestone	46.0	22 Sep 1999	P301	45	7682907965
2349	Bluestone	46.0	19 Oct 2001	S300	46	7629920821
2349	Bluestone	46.0	24 Sep 2003	P301	46	7682907965
2349	Bluestone	46.0	8 Nov 2005	P528	45	7609865463
2349	Bluestone	46.0	24 Nov 2007	P528	48	7609865463
4892	Bluestone	32.5	22 Sep 1999	S300	36	7629920821
4892	Bluestone	32.5	30 Sep 2001	S300	36	7629920821
4892	Bluestone	32.5	18Oct 2003	P301	37	7682907965

(question continues on next page)

A. Identify the Functional Dependencies in this table.

[3 marks]

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

[6 marks]

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

[6 marks]

D. Consider the following table, which records the examination results for students who have taken a particular subject. Students who fail a subject are not allowed to resit it.

STUDENT-RESULTS

SNUM	SUBJECT	EXAMDATE	MARK
454	History I	23-05-08	65
454	French I	17-06-08	74
538	History I	23-05-08	52
654	Psychology II	12-06-07	57

The primary key of this table is STUDENT (SNUM) + SUBJECT .

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

(a) STUDENT alone?

(b) STUDENT + SUBJECT+ EXAMDATE?

[4 marks]

(2) Suppose it is decided to allow students to re-sit exams in a subject which they have failed, in a following year. Would we need to change the definition of the primary key? Explain your answer.

[1 mark]

.

E. Consider the following table, which records the books and articles written by lecturers. (Lecturers are identified by Employee numbers, books by ISBNs, and articles by Serial Item and Contribution Identifiers (SICI). Note that the lecturer whose employee number is P22234 has written one book and two articles, and the lecturer whose employee number is K39423 has written two books and one article.

(question continues on next page)

LECTURER-DETAILS

Lecturer	Book	Article
P22234	1 84195 525 6	1046-8188(199501)13:1<69:FTTHBI>2.0.TX;2-4
P22234	.null.	0002-8231(199601)47:1<23:TDOMII>2.0.TX;2-2
K39423	0 86104 068 6	0095-4403(199502/03)21:3<12:WATIIB>2.0.TX;2-J
K39423	978 1 84489 416 1	.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]

[Total 25 marks]

Question 3

Consider a database created via the following statements.

```
CREATE TABLE Therapist (
        TherapistID    CHAR(8)
        Surname            VARCHAR(24),
        FirstName          VARCHAR(24),
        Department         VARCHAR(18),
        TherapyCode        CHAR(5),
        PRIMARY KEY        (TherapistID),
        FOREIGN KEY        (TherapyCode) REFERENCES Therapy (TherapyCode)) ;

CREATE TABLE Therapy (
        TherapyCode        CHAR(5),
        Description        VARCHAR(60),
        HourlyRate         INT,            -- in Dollars
        PRIMARY KEY        (TherapyCode)  );

CREATE TABLE Clinic (
        ClinicNum          INT,
        Start_date         DATE,
        End_date           DATE,
        Budget             INT,
        Manager            CHAR(8)
        PRIMARY KEY        (ClinicNum),
        FOREIGN KEY        (Manager) REFERENCES Therapist(TherapistID)) ;

CREATE TABLE Attendance (
        TherapistID        CHAR(8)
        ClinicNum          INT,
        AttendanceDate     DATE,
        HoursWorked        INT,
        PRIMARY KEY        (TherapistID, ClinicNum, AttendanceDate),
        FOREIGN KEY        (TherapistID) REFERENCES Therapist(TherapistID),
        FOREIGN KEY        (ClinicNum) REFERENCES Clinic(ClinicNum) );
```

THERAPIST

TherapistID	Surname	Department	TherapyCode

Therapy

TherapyCode	Description	HourlyRate

Clinic

ClinicNum	Start_date	End_date	Budget	Manager

Attendance

TherapistID	ClinicNum	AttendanceDate	HoursWorked

(question continues on next page)

A. Express the following natural language queries in SQL.

(1) List the codes and descriptions of all therapies with an hourly rate greater than 60 Dollars per hour, in alphabetical order of description. **[1 mark]**

(2) List the TherapistID and Surname of all Therapists who can administer the therapy described as 'Spinal Manipulation' who work in the 'Special Clinics' department. **[2 marks]**

(3) How many Therapists have the skill 'Spinal Manipulation'?. **[2 marks]**

(4) List all Clinic numbers and the total numbers of hours worked by therapists at that clinic, for Clinics where the total time worked by therapists has been less than 1600 hours. **[2 marks]**

(5) For all the Clinics 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 Therapist Surnames, Clinic numbers, date and number of HoursWorked, ordered by Clinic number and within that by Therapist Surname and within that Surname by date, for all Therapists who worked in the Clinics. **[3 marks]**

(6) List all Therapists (Surname and Department) with an hourly rate greater than the average hourly 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 Therapist can be allocated to a Clinic, per day, is 8; the number of HoursWorked has to be positive. **[2 marks]**

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

(3) **(a)** If the TherapistID is changed for a Therapist who is the manager of a Clinic, this update must be propagated in the table Clinic.

(b) A Therapist cannot resign (delete the tuple from Therapist) if he is managing a Clinic (note that a Clinic will be deleted from the database if it is closed down). **[3 marks]**

(4) The hourly rate for 'spinal manipulation' should be greater than the average rate. **[3 marks]**

(5) The manager of every Clinic must be from the 'Management' department. **[3 marks]**

[Total 25 marks]

Question 4

A. A large five star hotel which includes a gourmet restaurant has decided to set up a small database which will hold information on its chefs. Chefs are identified by unique Employee Numbers, which are eight characters long. They have Surnames, Middle Names, and First Names, each of which are up to 24 characters long. Each chef must have a Surname and a First Name, but not necessarily a Middle Name. The hotel wants to record each chef's date of starting work, and his monthly base pay, which must be more than 5,000 a month. Each chef may (or may not) also have a single Specialty (such as 'Salads' or 'Desserts), whose name can take up to 18 characters.
The hotel also wants to record information on the special dishes which can be prepared by each chef, although not all chefs have such special dishes. It wants to record the name of the dish, which will take up to 36 characters; and its type, which must be one of 'M' or 'O', (meaning 'Main Dish' or 'Other'). A dish in the database must be the dish of a single chef; if that chef leaves, the information about his dishes is deleted. Note that two chefs may be able to prepare the same dish, such as 'Steak Diane', so that a dish name alone is not a unique identifier.
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]

B.

(1) In the database context, what is a transaction? Give an example of a simple transaction.

[4 marks]

(2) Briefly explain the ACID properties which it is desirable for all transactions to have.

[4 marks]

C. When a system failure occurs, we can identify five types of transactions, from the viewpoint of the relationship of their start and completion times with respect to the last checkpoint and the time the system failed. Describe each of these types of transaction, stating, in each case, the corresponding recovery action that a DBMS must take (a diagram may help your explanation).

[5 marks]

D. In the context of a database management system, what is meant by a "deadly embrace"?

[2 marks]

[Total 25 marks]

Question 5

A. In the context of a database management system, what do we mean by a "system catalogue" (sometimes called a "data dictionary")? What types of data/information are usually found in a system catalogue, and what is the purpose of each type?

[6 marks]

B. Briefly define, in the database context, the terms "Data Definition Language" and "Data Manipulation Language" and provide two different examples of SQL statements illustrating each.

[6 marks]

C. Briefly discuss the basic differences between the external schema, the conceptual schema and the internal schema of a database.

[6 marks]

D. Consider the following two situations, both of which involve Suppliers, Products, and Shops:

In the first situation, a Supplier supplies Products to a central wholesale market. Shops purchase these Products, without knowing who supplied them. On a given day a shop may purchase a particular type of product that was supplied by several different suppliers. For example, Supplier S1 may supply Products P1, P2 and P3. Supplier S2 may supply products P2, P3, and P4. Shop H1 may buy Products P1 and P2. Shop H2 may buy Products P2 and P3.

In the second situation, there are also Suppliers, Products and Shops. But here, Suppliers contract directly with Shops to supply them with specific Products. For example, Supplier S1 may supply Product P1 to Shop H1, and also to Shop H2. The same supplier may supply Product P2 to Shop H2. Supplier P2 may supply Product P1 to Shop H1 also, and may also supply Shop H1 with Product P2.

(1) Draw Entity/Relationship diagrams for both situations, showing how Suppliers, Products and Shops are related.

[3 marks]

(2) Draw up relational schemas which could hold the information which we want to record in both situations. Be sure to indicate the primary keys of all relations.

[4 marks]

[Total 25 marks]

END OF EXAMINATION