UNIVERSITY OF LONDON

GOLDSMITHS COLLEGE

Department of Computing

B. Sc. Examination 2013

IS52021A
Database Systems

RESIT PAPER

Duration: 3 hours

Date and time:

This paper is in two parts, Part A and Part B. There are a total of three questions in each part. ***You should answer TWO questions from Part A and TWO questions from Part B.*** *Your answers to Part A and Part B should be written in separate answer books.*

Full marks will be awarded for complete answers to a total of four questions, two from Part A and two from Part B. Each question carries 25 marks. The marks for each part of a question are indicated at the end of the part in [.] brackets.

There are 100 marks available on this paper.

No calculators should be used.

THIS PAPER MUST NOT BE REMOVED
FROM THE EXAMINATION ROOM

PART A

Question 1

Consider the following relational schema corresponding to the tables Emp, Works and Dept storing data about the employees that work in departments. An employee can work in more than one department; the 'pct_time' field of the Works relation shows the percentage of time that a given employee works in a given department. The attribute 'managerid' in table Dept is the id of an employee (always identified by 'eid' in the table Emp), that is the manager of the department identified by the attribute 'did'. It is known that two or more departments may have the same manager. All the primary keys in the three relations are underlined.

Emp (eid: integer, ename: string, age: integer, salary: integer)
Works (eid: integer, did: integer, pct_time: integer)
Dept (did: integer, budget: real, dname: string, managerid: integer)

Write the following queries in SQL:

(a) Print/select the names and ages of each employee who works in both the Hardware department and the Software department. [3]

(b) For each department with more than 20 full-time-equivalent employees (i.e., where the part-time and full-time employees add up to at least 20 full-time employees), print the 'did' together with the number of employees that work in that department. [3]

(c) Print the name of each employee whose salary exceeds the budget of all of the departments that he or she works in. [3]

(d) Find the managerid of managers who manage only departments with budgets greater than 1,000,000. [3]

(e) Find the ename of managers who manage the departments with the largest budget. [3]

(f) If a manager manages more than one department, he or she controls the sum of all the budgets for those departments. Find the managerids of managers who control more than 5,000,000. [3]

(g) Find the managerid of managers who control the largest amount. [4]

(h) Find the employees names and ages that earn the second highest salary in the department Hardware. [3]

Question 2

Assume the database containing the following tables:

Hotel	(hotelNo, hotelName, city)
Room	(roomNo, hotelNo, type, price)
Booking	(hotelNo, guestNo, dateFrom, dateTo, roomNo)
Guest	(guestNo, guestName, guestAddress)

where Hotel contains hotel details, Room contains room details of which price is per day, Booking contains details of the bookings performed by the guests whose information is stored in table Guest. All the attributes are self-explanatory. The primary keys are formed of the underlined attribute in each table.

a) You are required to: [13]

i. Create a view HotelData containing the hotel name and the names of the guests staying at the hotel.
ii. Create a view BookingOutToday containing the account for each guest at Lido Hotel (note that CURRENT_DATE provides the date of today in SQL). More precisely, the view should provide, for each guest, all the guest's details and the amount s/he has to pay.
iii. Give the user Manager full access to view HotelData, with the privilege to pass the access on to other users.
iv. Remove user Frank the right to read from the view HotelData.
v. Can the view BookingOutToday be updated? Briefly justify your answer.

b) Consider the following view defined on the above database schema: [12]

```
CREATE VIEW HotelBookingCount (hotelNo, bookingCount)
AS SELECT h.hotelNo, COUNT(*)
FROM Hotel h, Room r, Booking b
WHERE h.hotelNo = r.hotelNo AND r.roomNo = b.roomNo
GROUP BY h.hotelNo;
```

For each of the following queries, state whether the query is valid and for the valid ones show how each of the queries would be mapped onto a query on the underling base tables (that is, illustrate the concept of view resolution using these queries). For the queries which are not valid briefly justify your answers.

i. SELECT * FROM HotelBookingCount;
ii. SELECT hotelNo FROM HotelBookingCount WHERE hotelNo = 'H001';
iii. SELECT MIN(bookingCount) FROM HotelBookingCount;
iv. SELECT COUNT(*) FROM HotelBookingCount;
v. SELECT hotelNo FROM HotelBookingCount WHERE bookingCount > 1000;
vi. SELECT hotelNo FROM HotelBookingCount ORDER BY bookingCount;

Question 3

a) Write an itemised list of up to eight functions that should be provided by a modern full-scale multi-user DBMS. [8]

b) Suppose you are involved in the development of the conceptual design of a database for the Sales Department of a company. The operation of the Department can be described as follows.

> They have a file of products that they provide to their customers. Each type of product has a unique product number, as well as a description, a cost and a price. The number of the product in stock and the number allocated are updated regularly. When the number in stock decreases to the reorder level, the product is reordered in a pre-decided quantity. They have a file of customers. Each customer is given a unique customer number. This file also contains customer names that consist of their first and last names, and customer addresses composed of street, city and postcode and the customer telephone number. Each customer has a credit limit, which is used to validate their orders. A customer may place zero, one or more orders at a time, and an order is always placed by one customer alone. Each order is identified by a unique order number. Other information as to orders includes the date due, the total price, and the status, that is, an order may be outstanding, partially delivered, or fully delivered and invoiced. An order may involve one or more than one type of products, and a type of products may be involved in more than one order. For each product being ordered in an order, its quantity, total price, and status (i.e., outstanding, partially delivered, or fully delivered) are recorded and updated regularly.

Given this information, draw an entity-relationship diagram for this Sales Department, including all the attributes and all the constraints that can be expressed. [17]

PART B

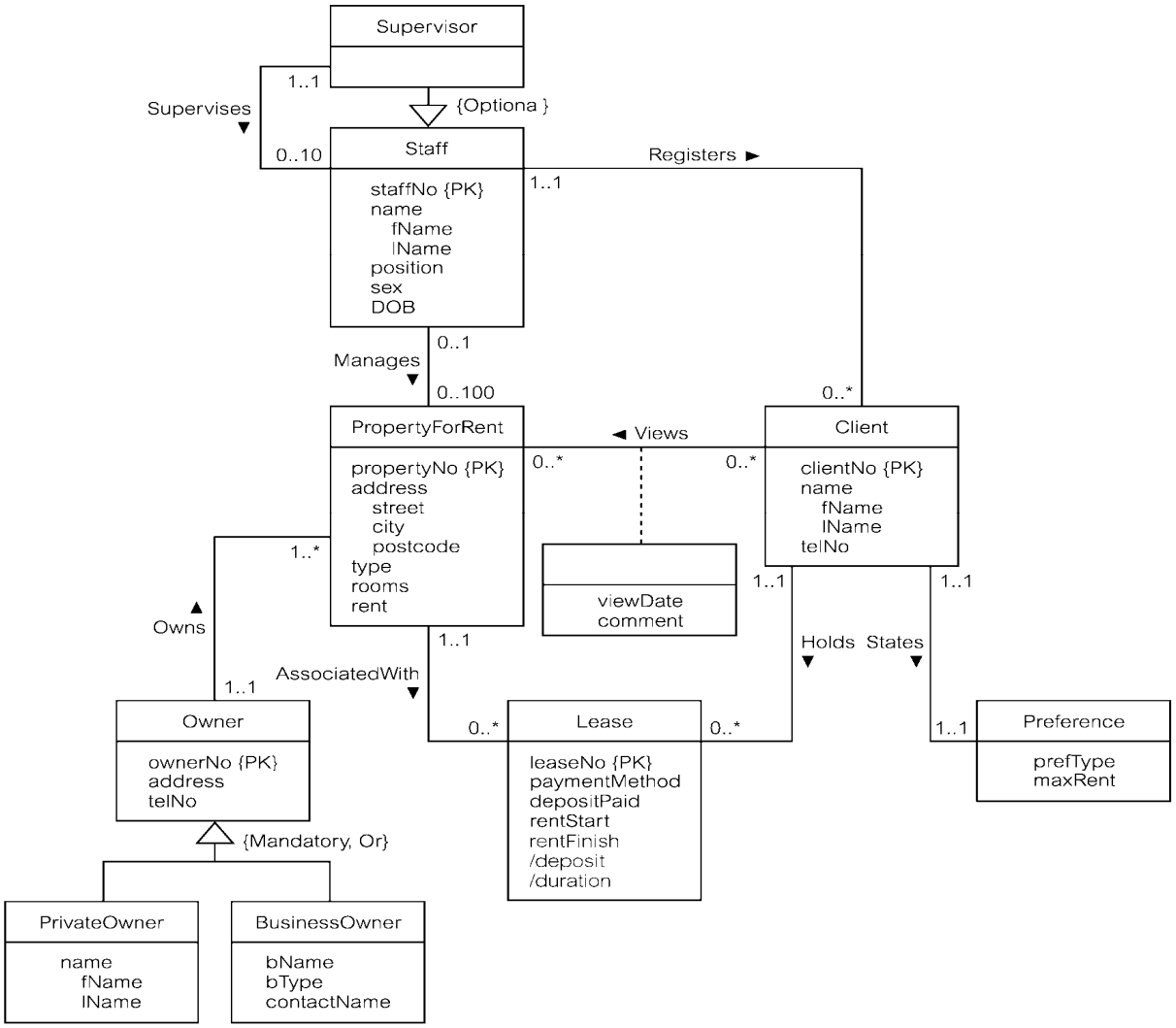
Question 4

In order to illustrate the logical and physical design of a database, do:

a) Convert the enhanced entity-relationship diagram below to the relational model. The tables should have clearly specified the attributes, the primary keys and the foreign keys. [19]

b) Choose a table generated at point (a) that has at least a foreign key and create it using SQL. [2]

c) Create an index that ensures a fast retrieval of data from the table from point (b). Provide the SQL code for two operations on this relation whose executions are speeded up and slowed down, respectively, because of this index. Explain your choice of operations in no more than two statements. [4]



Question 5

In the context of views in relational databases, do:

a) Briefly define view materialization and mention an advantage and a disadvantage of applying it (do not use more than five sentences for your answer). [5]
b) Provide an example of view that can be updated (the view should be expressed in SQL). Justify why the view is updateable. Illustrate with an example how an insertion is performed on this view. [8]
c) Provide three examples of views that cannot be updated (each view should be expressed in SQL, and each example should be of a different type). Justify in each case why the view is not updatable (use no more than two statements per case). [12]

Question 6

The table R below stores information about students that registered for various modules running during the Autumn and/or the Spring terms. The marks for the assignments in these modules are stored in the column "coursework" while the overall module marks are stored in the column "finalmark". Based on the final mark the results for modules consisting in a pass or a fail are stored in the column "result".

Studentid	Module	term	coursework	finalmark	Result
S001	CIS305	Autumn	70	60	Pass
S001	CIS305	Spring	50	60	Pass
S002	CIS305	Autumn	75	70	Pass
S002	CIS305	Spring	65	70	Pass
S001	CIS308	Autumn	36	33	Fail
S001	CIS308	Spring	30	33	Fail
S001	CIS312	Autumn	69	70	Pass
S001	CIS312	Spring	71	70	Pass

Assume the following functional dependencies:

fd1: (studentid, module, term) → coursework
fd2: (studentid, module) → finalmark
fd3: finalmark → result

a) Show that the functional dependencies fd1, fd2, fd3 are satisfied by the relation R. [6]

b) Check if the functional dependencies fd1, fd2, fd3 are left irreducible.
Justify your answers. [7]

c) Using the functional dependencies fd1, fd2, fd3, show that R is not in the Boyce-Codd normal form (BCNF), and then normalise R. Clearly indicate the candidate keys of each intermediary and final relations obtained by the application of Heath's theorem, and justify that the final tables are in BCNF. Note that for the normalisation of R you do not need to use the tuples but just the attributes of R or of any intermediary relations. [12]