Queen Mary
University of London

Wednesday 11th May 2016-2017   14:30 to 17:00

ECS519U/ECS740P   **Databases**   Duration: $2\frac{1}{2}$ hours

YOU ARE NOT PERMITTED TO READ THE CONTENTS OF THIS QUESTION PAPER UNTIL INSTRUCTED TO DO SO BY AN INVIGILATOR.

Instructions: This paper contains FOUR questions. **Answer ALL questions**.
Cross out any answers that you do not wish to be marked.

Calculators are not permitted in this examination.

Complete all rough workings in the answer book and cross through any work that is not to be assessed.

Examiners: Thomas Roelleke & Tony Stockman

Question 1

 Database (DB) History and Design

(a) DB history: The relational model was a major milestone in DB history.

  (i) Explain what type of model(s) the relational model replaced and explain why the relational model was seen as being superior to the model it replaced. Relate your answer to XML.

  (ii) Mention some of the challenges the relational model faced in the early and mid 90s.

[4 marks]

(b) DB design:

List the main steps of DB design. Explain each step briefly (one to three sentences).

[8 marks]

(c) SQL is the main language associated with the relational model. A SQL query is of the form

```
1  SELECT column_name, ...
2  FROM relation, ...
3  WHERE condition
```

Show and explain the relational algebra expression that corresponds to the SQL query.

[3 marks]

(d) Object-oriented software design involves concepts such as inheritance, generalisation, and specialisation.

  (i) Explain how the concepts are reflected in the extended entity-relationship model (E-ERM). Provide an example.

  (ii) Explain the process of specialisation, and the process of generalisation. Hint: Refer to the difference between entity types.

  (iii) List the four constraints of specialisation, and explain each constraint briefly.

  (iv) Explain the main two mapping options for mapping the E-ERM to the logical (i.e. relational) model.

[10 marks]

Question 2

Database Design and Entity-Relationship Model (ERM)

You are a DB consultant and your customer is the Eurostar company.

The Eurostar company serves the train lines between London and Paris, and between London and Brussels. The management asks you to propose a DB design.

The Eurostar trains run from London to Paris (or Brussels). There are several routes to be considered. For example, London St Pancras to Paris, and London Stratford to Paris. Also, some trains will terminate in Brussels, whereas others may continue to Cologne in Germany. Therefore, the DB design has to be ready to model train routes with different stations to start from, and different stations to terminate at.

The routes have one start station, several stations in the middle (referred to as stops), and one terminal station.

For each route, distance and journey time are to be stored. For example, the 306-mile route London-Paris takes 2hrs 50 mins. The max speed is 200 mph.

The company operates 20 trains. On average, 16 are in operation, and the others are in service (regular service checks and repairs). Trains can serve different routes.

The company employs around 400 people with many different roles. For example, there are the train crews operating the trains. A crew consists of two drivers, two conductors (one is the head of the crew), five service team members (responsible for drinks and food), and two security guards. Other employees are responsible for management and sales.

The DB system shall keep track of each train trip and the employees involved in the trip. Also, information about number and type of passengers are to stored for each trip. Then, queries such as "how many students travelled from Waterloo to Paris in 2016" can be processed.

 (a) (i) Design an entity-relationship model for the scenario described. Structure your design according to the design steps. Execute the design steps one, two and three, i.e. step number four is not required. Name and separate the design steps clearly.

   (ii) Provide a graphical representation of the ER model.

[15 marks]

 (b) Describe the main steps to map an ER model to the relational model. How do cardinalities affect the mapping?

[5 marks]

 (c) Explain the referential integrity constraint and provide an example (relations, attributes and tuples).

[5 marks]

Question 3

SQL

Given the following schema of a database:

```
1   # Logical model:
2   estate_agent(Id, Name, PhoneNumber);
3    office (Id, AgentId, Area, District , Sector, Unit, PhoneNumber);

5   property(Id, Area, District , Sector, Unit,
6           City, Price, Type, NumBedrooms, Description);
7   # Type: house or flat
8   # Description: a text  attribute

10  customer(Id, Name, PhoneNumber);

12  forSale(PropertyId, OfficeId) ;

14  forRent(PropertyId, OfficeId) ;

16  sold(PropertyId, OfficeId , CustomerId);


19  # To understand the semantics of Area, District , Sector and Unit,
20  # we consider the respective part  of the conceptual model:
21  #
22  #  office (Postcode, PhoneNumber, OpeningHours);
23  # Postcode is a composite attribute .
24  # postcode = (Area, District , Sector, Unit)
25  # For example:
26  # E1 4NS becomes
27  # Area="E", District ="1", Sector="4", Unit="NS"
```

Provide SQL expressions for the following queries. In all queries, make the join conditions an explicit part of the WHERE clause. Consider using intermediate views (or nested queries) where the query formulation becomes complex.

(a) Show the names of agents and the phone numbers of the offices that have 2-bedroom flats for sale. Post code of properties: E1.

[5 marks]

(b) Show the average of the prices for 2-bedroom flats for sale by estate agent "Fox".

[5 marks]

(c) Show all agents (name and phone number) who have 1-bedroom and 2-bedroom flats for sale in London.

[5 marks]

(d) Show the difference between the average price for 1-bedroom flats in post code E1 and post code W1.

[5 marks]

(e) Show all agents (name and phone number) who have 1-bedroom flats for sale and for rent.

[5 marks]

Question 4

  Transaction Management, DB Connectivity, DWH

(a) Database technology is associated with online transaction-based processing (OLTP) and with online analytical processing (OLAP).

  (i) For OLTP, the main task is to insure the consistency of data.
    List the main problems that may occur in transaction-based processing. An explanation is not required.

  (ii) Locking is one of the main technique to avoid the problems. Which problem arises from locking? Explain the problem and mention strategies to deal with the problem.

[5 marks]

(b) Database connectivity in applications using a DBMS.

  SQL queries retrieve sets of tuples. In JDBC and other API's, the result set is associated with a "handle".

  (i) Explain the design and main methods associated with the handle.

  (ii) What are the advantages/disadvantages of this design over an API that returns an array of tuples?

[5 marks]

(c) Explain the terms "copyright", "intellectual property" and "trademark".

[3 marks]

(d) Datawarehouse (DWH) technology:

  (i) DWH technology (i.e. OLAP) is different from OLTP. List five characteristics that explain the difference between OLAP and OLTP.

  (ii) Explain the schema of a DWH. In your answer, refer to two schemas, and explain the difference between the two.

  (iii) Why do IT professionals recommend to distinguish between OLAP and OLTP systems?

[12 marks]

End of questions