Two hours

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Databases

Date: Friday 28th January 2011
Time: 09:45 - 11:45

Please answer any THREE Questions from the FIVE Questions provided

For full marks your answers should be concise as well as accurate.
Marks will be awarded for reasoning and method as well as being correct.

This is a CLOSED book examination

The use of electronic calculators is NOT permitted
Question 1

(a) In the context of the relational model, what is the difference between a super key and a candidate key in a relation? Use an example to illustrate your answer. (4 marks)

(b) Consider the following relational schema from a database containing information about Employees and Projects:

   EmpProj(Employee#, Project#, Emp_Name, Proj_Name, Hours_Worked)

Employee# and Project# are unique for a given employee and project respectively. Does the above relational database design suffer from the problem of update anomalies? Explain your answer and provide an example. (4 marks)

(c) An operational database stores detailed data needed to support the day to day business processes and operations of a company. Explain the concept of data warehouses and differences between them and operational databases. Explain different types of applications that use data stored in data warehouses. (6 marks)

(d) Discuss the need for and advantages and disadvantages of object-oriented databases. Support your views with a discussion on emerging developments in the area and with a comparison to relational databases. (6 marks)
Question 2

(a) Describe the goal of the ANSI/SPARC three-level schema architecture of a database management system. Discuss how the three-level schema architecture can be applied to control access to information stored in a database system. (4 marks)

(b) You have been asked to design a database for an Art Museum Information System that satisfies the following requirements:

- The museum has a collection of ART_OBJECTs. We store each art object’s unique identification number, the year when it was created, a brief description of the work, the main title of the work (e.g., “Mona Lisa”), and a set of famous alternative titles for the work (e.g., “Gioconda”, “Smiling Lady”, etc.).
- An ARTIST creates one or more art objects. We store information about the duration of the creation process (i.e., how long the artist took to create the art object).
- The museum keeps track of artist’s name, date of birth, date of death, and provenance (country, region, city). The name is assumed to be unique.
- Art objects are shown at EXHIBITIONs. Each exhibition has a unique name, start date, end date, and duration derived from the start and end dates.
- Art objects are categorized (specialized) based on their type. There are three main types: PAINTING, SCULPTURE, STATUE, plus another type called OTHER to accommodate objects that do not fall into one of the three main types.
- A painting has a paint type (oil, watercolour, etc.), material on which the painting is drawn on (paper, canvas, wood, etc.), and style (modern, abstract, etc.)
- A sculpture has a material from which it was created (wood, stone, etc.), and also information regarding the height, weight and style
- For the museum, a statue is a sculpture that also records information regarding the inspiring theme (e.g., politician, emperor, scientist, soldier, etc.)

(i) Draw an ER diagram for the Art Museum Information System. Make sure that you include in your diagram all the requirements described above and that you specify the primary keys, cardinality constraints, participation constraints, and constraints on generalisation/specialisation. Please make sure that you also write down any additional assumptions that you make. (12 marks)

(ii) Provide an SQL command to create the EXHIBITION table. (4 marks)
Question 3

(a) Map the ER diagram shown below into a relational schema. Indicate primary key and foreign keys for each relation and show the dependencies between relations.

(b) When mapping an ER diagram to a relational schema, there are at least four options to map a generalisation/specialisation hierarchy. Explain the option that produces a single relation with one type attribute.

(c) Describe in plain English the purpose and execution effect of the following JDBC code:

```java
ResultSet rset = stmt.executeQuery("select branch_name, sum(balance) from account group by branch_name");
while (rset.next()) { System.out.println(rset.getString(1) + " "+ rset.getFloat(2)); }
```

(12 marks)
Question 4

(a) Explain the notion of referential integrity in the relational data model. Discuss different referential actions that can be specified within ON UPDATE and ON DELETE sub-clauses when defining a table. (6 marks)

(b) Consider the following relational schema for an airline database system

<table>
<thead>
<tr>
<th>Relation</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>FLIGHTS</td>
<td>flight-num, departure-city, destination-city</td>
</tr>
<tr>
<td>DEPARTURES</td>
<td>flight-num, flight-date, plane-type</td>
</tr>
<tr>
<td>PASSENGERS</td>
<td>passenger-id, name, address, preferred-meal, preferred-seat-location</td>
</tr>
<tr>
<td>BOOKINGS</td>
<td>passenger-id, flight-num, flight-date, seat-number</td>
</tr>
</tbody>
</table>

Underlined attributes in a relation are primary keys, whereas attributes with the same names in different relations are foreign keys. Provide SQL expressions for the following query/update requests:

- Find the names of all passengers preferring “aisle” seats.
- Count the number of passengers in the database.
- Create a view MAN747FLIGHTS consisting of the flight number, flight date and destination city of all flights departing from “Manchester” in a plane of type “747”.
- Find the flight numbers and flight dates where the number of booked seats is less than 50.
- Delete all bookings for the passenger with passenger-id = “pa121”. (10 marks)

(c) Describe situations where you may want to use a denormalised relational schema. Provide examples to illustrate your answer. (4 marks)
Question 5

(a) Consider the relational database described below. Explain whether the database violates inherent constraints and/or integrity constraints associated with relational databases.

<table>
<thead>
<tr>
<th>EmployeeNumber</th>
<th>Name</th>
<th>Salary</th>
<th>Dept_Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>5211</td>
<td>John Smith</td>
<td>100000</td>
<td>10</td>
</tr>
<tr>
<td>5212</td>
<td>Bob Jones</td>
<td>100000</td>
<td>30</td>
</tr>
<tr>
<td>5213</td>
<td>Mary Moon</td>
<td>125000</td>
<td>20</td>
</tr>
<tr>
<td>5214</td>
<td>John Smith</td>
<td>100000</td>
<td>10</td>
</tr>
</tbody>
</table>

Department

<table>
<thead>
<tr>
<th>Dept_Number</th>
<th>Dept_Name</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Sales</td>
<td>[Manchester, Cambridge]</td>
</tr>
<tr>
<td>20</td>
<td>Research</td>
<td>Surrey</td>
</tr>
</tbody>
</table>

(b) The following ODL (Object Definition Language) definitions form a partial model of an object database:

```object_definition_language
class Module
  (extent modules key code)
  {
    attribute string name;
    attribute string code;
    relationship <Academic> is_taught_by inverse Academic::teaches;
  }

class Academic
  (extent academics key staff_number)
  {
    attribute string name;
    attribute int staff_number;
    attribute string school;
    relationship set<Module> teaches inverse Module::is_taught_by;
  }
```

(i) Using the ODL, define an example class `Timetable` that contains three attributes (`day`, `time` and `room`) and one relationship (`lecture`, referring to the module taught in a given slot). Define an inverse relation as appropriate. (6 marks)

(ii) Using the OQL (Object Query Language), define a view `CS_timetable(day)` that retrieves the timetable for the given day (`day`) for all slots taught by a staff from the School of Computer Science. (3 marks)

(iii) Write an OQL query that retrieves a list of all academics from the School of Computer Science (ordered by their names) who teach all their modules in the same room. (5 marks)