ADVANCED DATABASE MANAGEMENT SYSTEMS

Tuesday 18th April 2023 - Afternoon

Answer any THREE questions out of FIVE. All questions carry equal marks.

Time: THREE hours

Answer any Section A questions you attempt in Answer Book A
Answer any Section B questions you attempt in Answer Book B

The marks given in brackets are indicative of the weight given to each part of the question.

Calculators are NOT allowed in this examination.
Section A

Answer Section A questions in Answer Book A

A1.

This question relates to the theoretical foundations of the Relational model and the emergence of alternative database models and supporting technology.

a) Compare and contrast the main features of Relational Calculus with those of Relational Algebra.

(6 marks)

b) Explain the main differences between each of the following data models:

i) Relational;

(4 marks)

ii) Object-Oriented;

(4 marks)

iii) Document Oriented.

(4 marks)

Your answer should cover differences in data representation, data definition languages, data manipulation, and data access/query languages.

c) Define the term “Data Persistence” and explain why it is important in a Database Management System (DBMS).

(7 marks)

A2.

Consider the following initial query tree (Fig A2.1) that operates on Tables EMPLOYEE (Fig A2.2), WORKS_ON (Fig A2.3), and PROJECT (Fig A2.4).

Fig A2.1 Query Tree

A database audit has shown that a staff member who changed division in a company has been able to access data that should not be accessed from their new division. Explain the concepts of authentication and authorisation, describe what they are and their relevance to the prevention of the incident.

(6 marks)

d) A database security administrator has many responsibilities, ranging from establishing abstract security policies to implementing security privileges for users. Consider the following tools available to a database security administrator and describe what they do:

i) SQL statement
   GRANT SELECT ON Payroll TO John

(2 marks)

ii) SQL statement
   REVOKE INSERT FROM HR_staff

(2 marks)

iii) SQL statement
   GRANT UPDATE ON product TO Amy WITH GRANT OPTION

(2 marks)

iv) Views.

(2 marks)

END OF EXAMINATION
Section B
Answer Section B questions in Answer Book B

B4.

a) The ACID properties of transactions are a standard set that ensures that database transactions are processed reliably.

i) State the name of each of the four properties and explain what it ensures. 

ii) Provide an example showing how simple transactions can lead to an undesirable outcome without adhering to the concept of atomicity.

b) Consider transaction schedules:

i) Explain the concepts of serial and serialisability when applied to transaction schedules.

ii) Provide an example of a set of database instructions (operations) performed by different processes that are not serialisable.

c) When processing transactions in distributed relational database systems, additional effort is required. Describe the two-phase commit protocol on the example of a transaction involving 3 RDBMSs on different servers by discussing the case where one RDBMS needs to abort its changes to local data. You may wish to use a diagram.

B5.

a) Consider data masking, data redaction and encryption:

i) Describe data masking.

ii) Describe data redaction.

iii) Explain how data masking can be more beneficial for developers than encryption during the testing phase of the development of new applications.

b) Database auditing has recently gained importance due to legislation such as GDPR (General Data Protection Regulation). What is database auditing, and how may it be conducted?
A3.

a) Describe each of the following approaches to distribution of data in a distributed database and comment on the benefits and drawbacks of each approach compared with storing data in a centralised database:

i) Data Fragmentation; (4 marks)

ii) Replication. (4 marks)

b) Company XYZ has three departments at different sites. Department 1 is located at Hull, department 2 is located at Manchester and department 3 is located at Boston.

The following Tables show sample data contained in the XYZ centralised database.

The database holds data on departments, employees and the assignments of employees to projects.

**Fig A3.1 EMPLOYEE (sample data)**

<table>
<thead>
<tr>
<th>EmployeeID</th>
<th>EmpName</th>
<th>Birthdate</th>
<th>Salary</th>
<th>Gender</th>
<th>Contact</th>
<th>DeptID</th>
</tr>
</thead>
<tbody>
<tr>
<td>12329</td>
<td>Rogers</td>
<td>1957-12-31</td>
<td>22480</td>
<td>M</td>
<td>12231</td>
<td>1</td>
</tr>
<tr>
<td>13453</td>
<td>Ahmed</td>
<td>1978-03-04</td>
<td>34000</td>
<td>F</td>
<td>13299</td>
<td>3</td>
</tr>
<tr>
<td>15463</td>
<td>Kohm</td>
<td>1956-06-03</td>
<td>29000</td>
<td>M</td>
<td>055456</td>
<td>3</td>
</tr>
<tr>
<td>16721</td>
<td>Collins</td>
<td>1959-04-23</td>
<td>45000</td>
<td>F</td>
<td>454433</td>
<td>2</td>
</tr>
<tr>
<td>16722</td>
<td>Sakar</td>
<td>1979-05-27</td>
<td>49000</td>
<td>M</td>
<td>563476</td>
<td>2</td>
</tr>
<tr>
<td>16779</td>
<td>Mehmet</td>
<td>1969-03-04</td>
<td>55000</td>
<td>F</td>
<td>073037</td>
<td>1</td>
</tr>
</tbody>
</table>

**Fig A3.2 ASSIGNMENT (sample data)**

<table>
<thead>
<tr>
<th>EmployeeID</th>
<th>ProjectID</th>
<th>TotalHrsWorked</th>
</tr>
</thead>
<tbody>
<tr>
<td>13453</td>
<td>43</td>
<td>200</td>
</tr>
<tr>
<td>15463</td>
<td>46</td>
<td>19</td>
</tr>
<tr>
<td>16721</td>
<td>47</td>
<td>22</td>
</tr>
<tr>
<td>16722</td>
<td>47</td>
<td>22</td>
</tr>
<tr>
<td>16779</td>
<td>49</td>
<td>450</td>
</tr>
<tr>
<td>12329</td>
<td>49</td>
<td>450</td>
</tr>
</tbody>
</table>

**Fig A3.3 PROJECT (sample data)**

<table>
<thead>
<tr>
<th>ProjectID</th>
<th>ProjectName</th>
<th>Projectstarted</th>
<th>DeptID</th>
</tr>
</thead>
<tbody>
<tr>
<td>43</td>
<td>Aquarius</td>
<td>2021-02-03</td>
<td>3</td>
</tr>
<tr>
<td>46</td>
<td>Markets</td>
<td>2020-12-11</td>
<td>3</td>
</tr>
<tr>
<td>47</td>
<td>Hospitality</td>
<td>2020-12-24</td>
<td>3</td>
</tr>
<tr>
<td>49</td>
<td>Networks</td>
<td>2016-11-21</td>
<td>2</td>
</tr>
<tr>
<td>50</td>
<td>Catering</td>
<td>2015-04-06</td>
<td>1</td>
</tr>
</tbody>
</table>

**Fig A3.4 DEPARTMENT (sample data)**

<table>
<thead>
<tr>
<th>DeptID</th>
<th>DeptName</th>
<th>Manager</th>
<th>DepartmentLocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>R&amp;D</td>
<td>16722</td>
<td>Boston</td>
</tr>
<tr>
<td>2</td>
<td>Admin</td>
<td>16721</td>
<td>Manchester</td>
</tr>
<tr>
<td>1</td>
<td>HQ</td>
<td>16779</td>
<td>Hull</td>
</tr>
</tbody>
</table>

Users and applications in the Boston and Manchester departments require fast access to each of the four tables to process information only for the department that is located at that site. The attributes they actually use from the EMPLOYEE data are Empname, EmployeeID, Salary and DeptID.

The company headquarters are located in the department at Hull and users at that site access all information about employees, departments, assignments and projects frequently.

With reference to the scenario and the sample tables, produce an appropriate fragment allocation design. Justify your answer. (8 marks)

c) Over recent years, mobile computing devices (e.g., smartphones, wearable devices and PDAs) have been widely used to store and share data captured on a mobile device connected over a mobile network to a centralised database. Mobile devices use a “lightweight” database embedded in the mobile device, and numerous apps provide an interface for user access to that data.

Describe a simple mobile application you are familiar with and briefly explain the key issues of data distribution and how data consistency is maintained, for example, when communication with the central database fails. (9 marks)