ECM2419

UNIVERSITY OF EXETER

COLLEGE OF ENGINEERING, MATHEMATICS AND PHYSICAL SCIENCES

COMPUTER SCIENCE

Examination, January 2022

Database Theory and Design

Module Leader: Dr Zeliang Wang

Duration: TWO HOURS

No word count

Answer ALL questions.

The marks for this module are calculated from 60% of the percentage mark for this paper plus 40% of the percentage mark for associated coursework.

This is a CLOSED BOOK examination.

Question 1

(a) A large organisation has adopted a database system rather than a file based system. Briefly explain five advantages of such decision.

(10 marks)

(b) Consider the following relations that form part of a database, where the primary keys are underlined. Note that the relation *Employee* contains employee details; the relation *Department* contains department details and the attribute *mgrEmpID* identifies the employee who is the manager of the department. There is only one manager for each department; the relation *Project* contains details of the projects in each department, and no two departments can run the same project; the relation *WorksOn* contains details of the hours worked by employees on each project.

Employee (empID, fName, lName, address, DOB, sex, position, deptNo)
Department (deptNo, deptName, mgrEmpID)
Project (projNo, projName, deptNo)
WorksOn (empID, projNo, hoursWorked)

Write the following queries in relational algebra.

(i) Find all the details of employees who are male.

(2 marks)

(ii) List the names and addresses of all employees who are Managers.

(4 marks)

(iii) Produce a list of the names and addresses of all employees who work for the 'HR' department.

(5 marks)

(iv) Produce a list of the names of all employees who work on the 'STF' project.

(5 marks)

(c) Provide two reasons that Strict Two-Phase Locking (2PL) is popularly used in many database systems.

(4 marks)

(d) If the number of *read-only transactions* is increased, what would happen to the database system throughput? **(3 marks)**

(e) What are the two types of database updating schemes in log-based recovery? Explain the difference between them.

(7 marks)

(Total 40 marks)

Question 2

(a) A GP practice in Exeter has a number of doctors, with each of whom a number of patients are registered. The practice wishes to record on a computer details of drugs prescribed for its patients, and also the dates and times of consultations between the patients and particular doctors in the practice. Prescriptions, which may be for one or more drugs, are given either in consultations, or as a repeat of prescriptions previously given and signed in a consultation. All repeat prescriptions, which are dated, must be signed by a doctor in the practice, not necessarily the same one who signed the original prescription. Prescriptions give not only the name of the drug, but also the amount and frequency of the dose, the total quantity to be supplied, and the date, if any, when a repeat prescription would be due.

Draw an *entity-relationship diagram* (using UML notation) for the patient management system as described above. Identify all entities, relationships, attributes (you may add some basic information for each entity), primary keys and multiplicity constraints.

(18 marks)

(b) The following table shows some sample data about mortgage application appointments from a bank. Each mortgage applicant is given an appointment at a specific date and time with a mortgage advisor at a particular branch of the bank. Assume that for each day of appointments, a mortgage advisor is allocated to a specific branch for that day. Each applicant is registered at only one branch, and may have more than one appointment on a given day.

date	time	staffNo	staffName	applicantNo	applicantName	branchNo
19/06/18	9:00	S022	Paul	A030	Josh	B11
19/06/18	11:00	S022	Paul	A035	Julian	B11
19/06/18	9:00	S028	John	A039	Fabian	B05
21/06/18	13:00	S028	John	A039	Fabian	B05
21/06/18	15:30	S010	Emma	A035	Julian	B11
22/06/18	17:00	S010	Emma	A050	Liz	B03

(i) The above table is susceptible to update anomalies. Provide an example of deletion anomaly.

(2 marks)

(ii) Identify a *candidate key* for the table.

(2 marks)

(iii) Transform the table into relations in the *2nd Normal Form* and then *3rd Normal Form* by using relational schemas (with primary keys underlined) rather than detailed tables.

(8 marks)

(Total 30 marks)

Question 3

Given the following database in which the primary keys are underlined. The relation *Catalog* lists the prices charged for *Parts* by the *Suppliers*. Write expressions in SQL for each of the following queries.

Suppliers (supplierID, supplierName, address)
Parts (partID, partName, color)
Catalog (supplierID, partID, cost)

(a) List the names of suppliers who supply a black part that costs less than 200.

(6 marks)

(b) Retrieve the names of parts supplied by *Smart Widgets* and no one else.

(6 marks)

(c) For each part, retrieve the name of the supplier who charges the most for that part.

(6 marks)

(d) Retrieve the IDs of suppliers who supply only black parts.

(6 marks)

(e) Retrieve the IDs of suppliers who supply at least a blue part and at least a yellow part.

(6 marks)

(Total 30 marks)