The Open University of Sri Lanka

Faculty of Engineering Technology

Department of Electrical and Computer Engineering

Study Programme	: Bachelor of Software Engineering Honours
Name of the Examination	: Final Examination
Course Code and Title	: **EEI5466/EEX5466/EEI5566–Advanced Database Systems**
Academic Year	: 2020/21
Date	: 21ˢᵗ December 2021
Time	: 1400-1700hrs
Duration	: **3 hours**

General Instructions

1. Read all instructions carefully before answering the questions.

2. This question paper consists of **Eight (8)** questions in **Nine (9)** pages. Answer **FIVE** questions **ONLY.**

3. Answer for each question should be written in the book provided and each question should commence from a new page.

4. This is a Closed Book Test **(CBT).**

5. Answers should be in clear handwriting.

6. Do not use red colour pen.

Question 1 – Database Programming   (20 marks)

Consider the following relations in a bank database:

Branch (_bName_:char(15), _city_:varchar(12), _phone_:varchar(10))
Customer (_custno_:integer, _cName_:char(15), _gender_:char(1), _birthdate_:date)
Account (_accno_:integer, _branchname_:char(15), _acctype_:char(1), _balance_:real)
AC (_accno_:integer, _custno_:integer)

The attributes of the Branch relation are name of the branch (_bName_), _city_, and _phone_. The Customer relation has attributes to record the customer number (_custno_), name (_cName_), _gender_ ('M' or 'F'), and _birthdate_ of customers. The Account relations consists of account number (_accno_), branch name (_branchname_), account type (_acctype_) which may be individual (acctype='I') or joint (acctype='J'), and account balance (_balance_). The attribute _branchname_ in Account is a foreign key that references Branch; an individual account belongs to a single customer, while a joint account is held by two or more customers. The attribute _accno_ and _custno_ in AC reference the relations Account and Customer respectively. The primary keys of all relations are underlined.

(a) Express the following queries in SQL.
  i. For each customer with a total balance exceeding 50000LKR for all his/her accounts together, display the customer number and total balance.

                                                                        (3 marks)

  ii. For each branch, display the branch name, account number and balance of each account that has a balance greater than twice the average balance of all accounts at that branch. In the resulting report, order the accounts of each branch in the descending sequence of account balance.
                                                                        (5 marks)

(b) It is required to record the number of account holders for each account by adding an attribute named _holders_. The value of this attribute will be 1 for individual accounts and greater than 1 for joint accounts.

  i. In the Account table, add a new attribute named _holders_ of integer type with a default value of 1.
                                                                        (1 marks)

  ii. Write a T-SQL procedure to update the holders attribute (that has just been added) only for joint accounts.

                                                                        (5 marks)

  iii. Write a trigger to decrease the holders attribute value when a row is deleted from the table AC only if it is for a joint account. If the updated value of holders is one for the joint account, then change the acctype to 'I'.

                                                                        (6 marks)

Question 2 – Object Relational Databases   (20 marks)

Consider the following object relational database schema for a database of insurance policies for cars:

Object types:

Customer_t (cid: char(6), name: varchar(15), birthdate: date, phone: char(10), address: varchar(50))
Car_t (regno: char(9), make: varchar(12), model: varchar(10), mdate: date, owner: ref customer_t, value: number(8,2))
Claim_t (claimno: char(12), cdate: date, amount: number(8,2), claimant: ref customer)
Claim_ntab table of claim_t
Policy_t (pid: char(7), sdate: date, edate: date, inscar: ref car_t, premium: number(6,2), claims: claim_ntab)

Tables:

Customers of Customer_t (cid primary key)
Cars of car_t (regno primary key, owner references Customers)
Policies of policy_t (pid primary key, inscar references Cars)
Nested table claims store as claims_ntable

The tables named Customers, Cars, and Policies contain tuples for all customers, cars and policies respectively. Their attributes are indicated by the corresponding types.

Attributes of Customers are customer id (cid), name, date of birth, phone and address. Attributes of Cars are registration number (regno), make, model, date of manufacture, owner of the car, and insured value of the car.

Attributes of Policies policy id (pid), starting date, ending date, insured car, annual premium and the claims made.

Attributes of the nested table of claims are claim number (claimno), claim date (cdate), the amount of claim and the claimant.

The attribute types are specified in the type descriptions above, as also are the primary keys and referential constraints in the table schema.

(a) Write Oracle OBJECT SQL statements to answer the following queries (use columns of REF type instead of joins to link tables):

(i) Find the average insurance premium on cars owned by customers aged between 20 and 25 years. Sysdate contains current date and the function month_between(d1,d2) gives the number of months in decimals between the dates d1 and d2 where d1>d2.

(3 marks)

(ii) For each make and model of car, find the total claim amount on policies that expire between 1 Jan 2020 and 31 Dec 2020.

(4 marks)

(b) Write Oracle object SQL to insert a claim against an existing policy that has a pid of SL12354, given the following claim details: claim number: 001, claim date: 12 July 2020, claim amount: 2000, and customer id of claimant: S25431. Assume that the claimant is already present as a customer in the database.

(3 marks)

(c) Write Oracle object SQL to calculate the renewal premium for a given policy in the database using the following logic:

If the policy had no claims or the total claim is less than 1000, then the new premium is the same as the current premium. If the total claim on a policy is greater than or equal to 1000, then the new premium is to be 20% more than the current.

(7 marks)

(d) Using the method defined above in (c), get the renewal premium for the policy of a car with registration number SLA984.

(3 marks)

Question 3 – Relational Algebra
(20 marks)

Consider the following relations containing airline flight information:

```
Flights  (flno, from, to, distance, depart, arrives)
Aircraft (aid, aname, travelRange)
Certified (eid, aid)
Employee (eid, ename, salary)
```

Note that the Employee relation includes pilots and other kinds of employees as well; every pilot is certified for some aircraft (otherwise, he or she would not qualify as a pilot), and only pilots are certified to fly.

Write the following queries in **relational algebra.**

(a) Find the names of the pilots certified for some Boeing (i.e. aname) aircraft.

(3 marks)

(b) Identify and print *eid* and *ename* of all the pilots whose salary is more than Rs. 100,000/=.

(3 marks)

(c) Find the *aid* of all aircraft that can be used on non-stop flights from *Madras* to *Colombo*. (Hint: consider *travelRange > distance* to identify non-stop flights)

(4 marks)

(d) Find *ename* of pilots who have been assigned more than three aircrafts.

(5 marks)

(e) Find *eid* and *ename* of all employees and print aircraft information (i.e. *aid*) for pilots.

(5 marks)

iv. What indexes would you suggest speeding up the query plan considered in (iii.)? Assume that DBMS supports only B+ tree indexes.

(3 marks)

Question 6 – Data Mining Applications and Association Rule Mining (20 marks)

(a) Discuss whether or not each of the following activities is a data mining task.

i. Grouping the customers of a company according to their gender.

(1 marks)

ii. Predicting the future stock price of a company using historical records.

(2 marks)

iii. Monitoring the heartbeat of a patient for abnormalities.

(2 marks)

(b) Suppose that you are employed as a data mining consultant for an Internet search engine company. Describe how three of the data mining techniques given below can help the company, by giving specific examples; clustering, classification, association rule mining, and anomaly detection.

(6 marks)

(c) What is the *Apriori principle*?

(2 marks)

(d) Consider the following transactional database.

TID	Items Bought
1	Milk, Beer, Diapers
2	Bread, Butter, Milk
3	Milk, Diapers, Bread
4	Bread, Butter, Cookies
5	Beer, Bread, Diapers
6	Milk, Diapers, Bread, Butter
7	Bread, Butter, Diapers
8	Beer, Diapers
9	Milk, Diapers, Bread, Butter
10	Beer, Cookies

i. Use the apriori principle and find frequent itemsets with minimum support of 70%.

(4 marks)

ii. Find all association rules with minimum support of 70% and minimum confidence of 50%.

(3 marks)

Question 7 – Transactions and Concurrency Control   (20 marks)

(a)   Briefly explain the properties of a transaction.

                                                                        (4 marks)

(b)   Briefly explain, what a *Serializable Schedule* is?

                                                                        (1 mark)

(c)   Briefly explain the rules in Strict 2 Phase Locking Protocol.

                                                                        (2 marks)

(d)   Consider a database with objects X and Y and assume that there are two transactions T1 and T2. Transaction T1 reads object X and Y and then writes object X. Transaction T2 reads objects X and Y and then writes objects X and Y. Both T1 and T2 commit after all read and write actions of them.

   i. Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-read conflict.

                                                                        (2 marks)

   ii. Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a read-write conflict.

                                                                        (2 marks)

       Give an example schedule with actions of transactions T1 and T2 on objects X and Y that results in a write-write conflict.

                                                                        (2 marks)

   iii. For each of the three schedules, show that Strict 2PL disallows the conflicts.

                                                                        (2 marks)

(e)   Consider the following sequence of actions, listed in the order they are submitted to the DBMS. The Strict 2PL has been used for concurrency control.

       T1:R(P), T2:W(Q), T2:W(P), T3:R(R), T3:R(Q), T1: W(R)

| T₁	T₂	T₃
S(P)		
R(P)		
	X(Q)	
	W(Q)	
	X(P)	
	W(P)	
		S(R)
		R(R)
		S(Q)
		R(Q)
X(R)		
W(R)		

Assume that older transaction has higher priority always. The DBMS processes actions in the order shown.

i. Follow wound-wait policy to deal with deadlock in above schedule. Draw schedule again.

(3 marks)

ii. Follow deadlock detection approach to deal with deadlocks in above schedules. Draw a wait-for graph for the schedule given above to detect the deadlock and briefly explain how to identify the deadlock and remove.

(2 marks)

Question 8 – Crash Recovery   (20 marks)

(a.) Consider the diagram below:
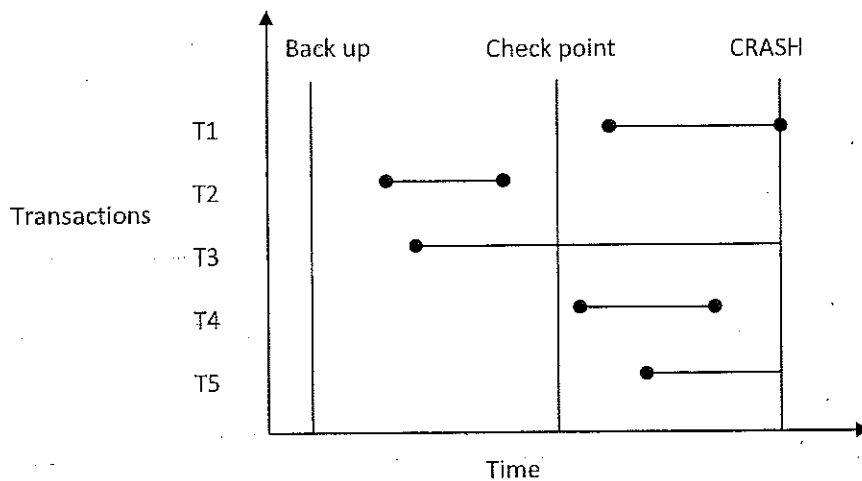
What is the desired state that the database should be in after Crash Recovery? Briefly explain.

(5 marks)

(b.) What is meant by Write-Ahead Logging. Why it is necessary to implement this protocol in a DBMS?

(5 marks)

(c.) What are the roles of Analysis, Redo and Undo phases in ARIES?

(5 marks)

(d.) What are the differences between update log records and CLRs?

(5 marks)