VIT-AP UNIVERSITY

	Final Assessment Test – Winter 2023-24 Semester - May 2024	
	Max Marks: **100**	Duration: **3 Hours**
Course Code: **CSE2007**	Course Title: **Database Management Systems**	
Question Paper ID: **009821**	Exam Type: **CLOSED BOOK**	School: **SCOPE**
Date: **09/05/2024**	Slot: **A1+TA1**	Session: **FN**

**Keeping mobile phone/smart watch, even in 'off' position is treated as exam malpractice.**

General Instructions if any:

1. "*fx* series" - non Programmable calculator is permitted: No
2. Reference tables permitted: No

Section - 1: Answer any 10 questions. (10 × 10 = 100 Marks)

		Marks	CO	BL

Q1. You have a database company. Your current project is to design a database for art galleries for capturing all the information galleries need to maintain. Galleries keep information about artists, their names (which are unique), birthplaces, age, and style of art. For each piece of artwork, the artist, the year it was made, its unique title, its type of art (e.g., painting, lithograph, sculpture, photograph), and its price must be stored. Artwork is also classified into groups of various kinds. A given art piece may belong to more than one group. Each group is identified by a name that describes it. Galleries keep information about customers. For each customer, galleries keep that person's unique name, address, and total amount of rupees spent buying the art from the gallery. Draw the ER diagram for the art gallery database.
10 1 4

Q2. Suppose you are designing a database for a library management system. Details about the books and library members are maintained by the library. Members can issue books. They have to return the books on time. Explain the different integrity constraints required with examples in the context of the library management system. Write down the roles of database administrator in the above database.
10 1 4

Q3. Convert the ER diagram to a relational model.
10 2 3
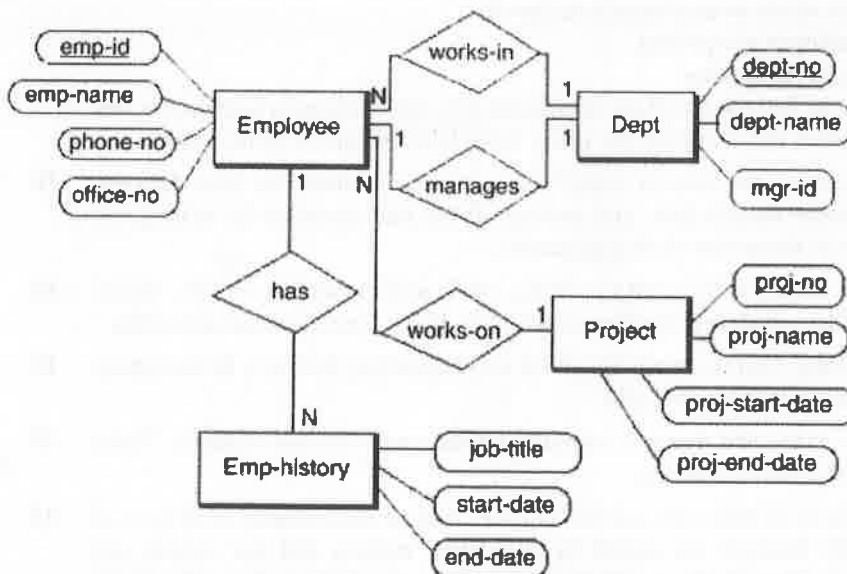
Q4. Consider the following relational database schema consisting of the four relation schemas: **10   2   4**
passenger (passanger_id, name, gender, city)
agency (agencyid, agency_name, agency_city, agency_state)
flight (flight_id, fdate, time, source, destination)
booking (passanger_id, agency_id, flight_id, fdate)
Answer the following questions using relational algebra queries.
a) Find the details of all female passengers who are associated with agency named MakeMyTrip.
b) Find the agency names who do not have any bookings for passenger with id 768.
c) Find the agency names that located in the same city as passenger with name Rahul.
d) Find the passenger names for those who do not have any bookings in any flights.

Q5. Consider a relation Student (StudentID, ModuleID, ModuleName, StudentName, StudentAddress,   **10   3   4**
TutorId, TutorName) with the following functional dependency.
StudentID → StudentName, StudentAddress, TutorId, TutorName
ModuleID → ModuleName
TutorID → TutorName
Normalize the above table upto 3rd Normal Form.

Q6. What are the properties for a decomposition to be lossless and dependency preserving?   **10   3   4**
Let R = {ssn, ename, pnumber, pname, plocation, hours} and R is decomposed into three relations
R1, R2, and R3 as follows;
R1 = EMP = {ssn, ename}
R2 = PROJ = {pnumber, pname, plocation}
R3 = WORKS_ON = {ssn, pnumber, hours}
Assume that the following functional dependencies are holding on relation R.
ssn → ename
pnumber → {pname, plocation}
{ssn, pnumber} → hours.
Find whether the decomposition into R1, R2, and R3 is lossless join decomposition or not.

Q7. Given the following relations   **10   4   4**
Employee(Eno, Ename, Contact)
Proj_Assigned(Eno, Project_No, Designation, PDuration)
Where,
Eno is Employee number,
Ename is Employee name,
Project_No is Project Number in which an employee is assigned,
Designation is the role of an employee in a project,
PDuration is duration of the project in months.
Write a relational algebra query to find the list of all employees who are working in a project as the
manager which is more than 2 years old. Optimize the query using heuristic query optimization.

Q8. What are the different join operation we have in SQL. Take a sample database for food delivery   **10   4   2**
system. Write the tables with some sample data and explain all the join operation by writing the
appropriate SQL query. Also write the output of all join queries.

Q9. For the following schedule S: r2(Y), w2(Y), r3(Y), r1(X), r4(X),w1(X), w4(Y), w3(Y), r2(X),   **10   5   4**
r1(Y), w1(Y), w4(X) Check if the schedule is conflict serializable. If yes Find the serial schedules.

Q10. What is the use of 2-Phase locking. Can we avoid Deadlock and Cascading Rollback In two phase   **10   5   1**
locking. Give your answer by using suitable example.

Q11. What is the difference between static and dynamic hashing. Explain with suitable example. Write   **10   6   1**
down the properties of static and dynamic hashing.

Q12. Suppose we have a database file of 45,000 rows and 10 columns. Size of each column is 20 byte. If   **10   6   3**
the disk block size is 1.25KB. Records are stored in unsapnned manner and the records are
unsorted. Secondary indexing is done on the search field which has all unique values. The block
pointer is 10 byte. What would be the number of disk accesses to do a successful lookup for a
record

End of Question Paper

Final Assessment Test – Winter 2023-24 Semester - May 2024	
Max Marks: **100**	Duration: **3 Hours**

Course Code: **CSE2007**	Course Title: **Database Management Systems**	
Question Paper ID: **009709**	Exam Type: **CLOSED BOOK**	School: **SCOPE**
Date: **10/05/2024**	Slot: **A2+TA2**	Session: **FN**

Keeping mobile phone/smart watch, even in 'off' position is treated as exam malpractice.

General Instructions if any:

1. "*fx* series" - non Programmable calculator is permitted: No
2. Reference tables permitted: No

Section - 1: Answer any 10 questions. (10 × 10 = 100 Marks)	Marks	CO	BL

Q1. Consider a system for an online marketplace where users can buy and sell handmade crafts. How does three schema architecture facilitate Data Independence in the database system? Provide concrete examples from both the business and technical perspectives to illustrate your answer. **10** / 1 / 2

Q2. Major airline companies that provide passenger services in Singapore are UniAir, TransAsia Airways, Far Eastern Transport, Great China Airlines, etc. **10** / 2 / 2

Singapore Federal Flight Administration (SFFA) keeps a database with lots of information on all airlines. The information consists of: Each airplane has an
identification number, name of the contact person, and telephone number. For each aircraft identification number, capacity and model are recorded.

Each employee has an employee identification number, name, address, birthday, sex, position with the company and qualification.

Each route has a route identification number, origin, destination, classification (into domestic or international route), distance of the route and price charged per passenger.

Each airline keeps the information about their buy/sell transactions (for example, selling an airplane ticket is a sell transaction, and paying for maintenance is a buy transaction).

Each transaction has a transaction identification number, date, description, and amount of money paid/received.

Design a suitable ER diagram for the above scenario with all possible entities, attributes and the corresponding relationship with its participations.

Q3. Considering the following Library Relational schema. Write the queries in relational algebra format. **10** / 2 / 4

    Book(Book_id, title, publisher)
    Book_Authors(Book_id, author_name)
    Book_loans(Book_id, Card_no, Date_out, Due_date)
    Borrower(Card_no, Name, address)

  a.   Find the books written by 'Ramakrishnan' or published by 'Tata Mcgraw hill'.
  b.   List the titles of the books borrowed by members whose names start with the letter 'A'.
  c.   Find the books published by 'pearson education' which are taken by a person named 'John'.
  d.   Find the members who have taken more than three books.
  e.   Find the card numbers for which the members have taken all the books of 'Wiley' Publications.

Q4. Map the following ER diagram to a relational model. Identify the relationship and map the cardinality. Represent primary key and foreign key in the schema. **10 2 2**
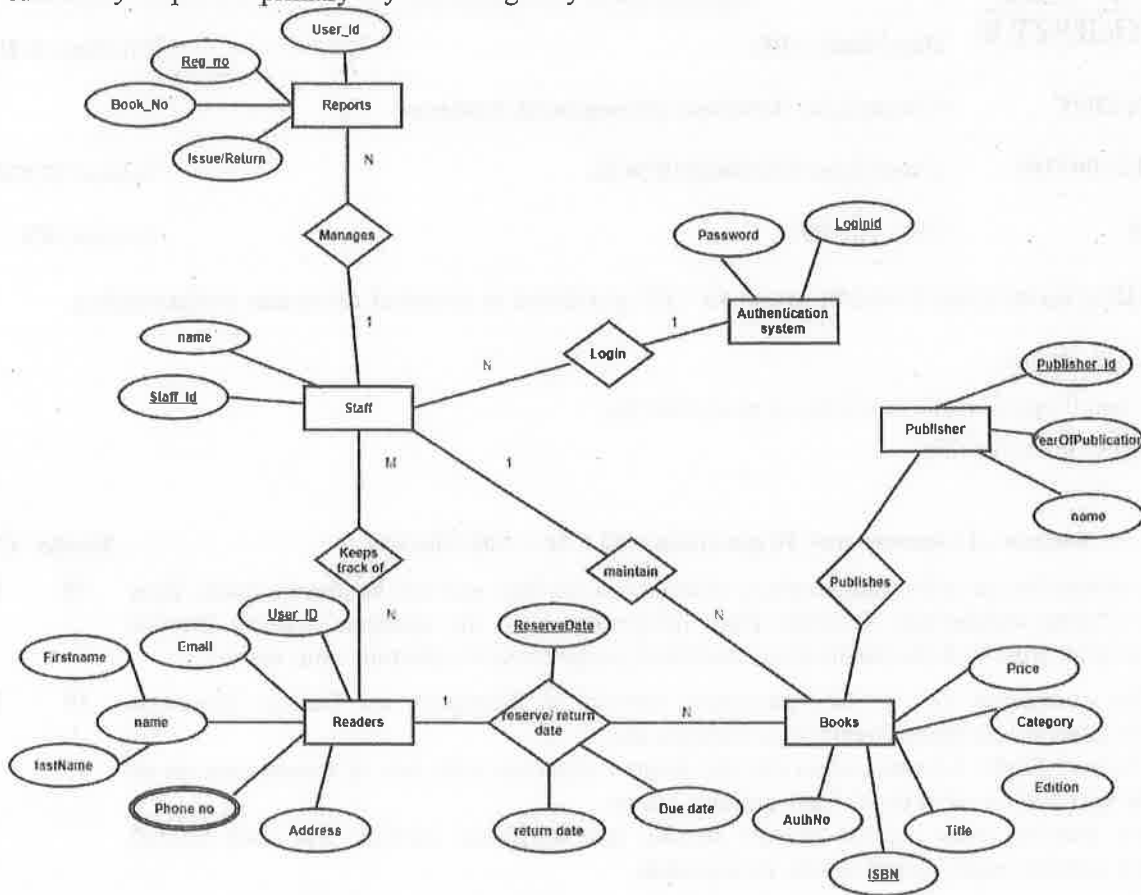
Q5. A relation R has attributes A, B, C, D, E, and satisfies the following FDs: **10 3 3**
A ->BC
B ->D
CD -> E
E ->A
a) Show that (A, B, C) and (A, D, E) are a lossless decomposition of R.
b) Show that (A, B, C) and (C, D, E) are not a lossless decomposition of R.

Q6. A) A relation schema R = (A, B, C, D, E) and for the following set of functional dependencies: **10 3 3**
$\quad$ F = {A → BC, B → C, A → B, AB → C }.
$\quad$ Compute the minimal cover for F.

B) Consider the following relation scheme R (A, B, C, D, E, F, G, H, I, J}
together with the following functional dependencies:
{F-> H, E),
{{C, G, A} -> {B, D, J}},
{{B}->{A, I}},
{B, G} ->{C}.
Use Armstrong's axioms to prove formally that this set of functional dependencies implies that {B, G} -> {D}.

Q7. Draw the initial query tree for the following query and show the steps to optimize the query. **10 4 3**
$\quad$ SELECT P.PNO, P.DUM, E.LNAME, E.ADDRESS, E.DOB FROM PROJECT P, DEPARTMENT D, EMPLOYEE E WHERE P.DUM=D.DNO AND D.MGRSSN=E.SSN AND P.LOCATION='HYDERABAD'.

Q8. Consider the relational database, where the primary keys are underlined. Give an expression in SQL   **10**   **4**   **4**
for each of the following queries:
Employee (employee_id, name, department_id, salary)
Department (department_id, name, location)
Project (project_id, name, department_id)
Works_on (employee_id, project_id, hours)

a) Display the project IDs and names for all projects
b) Find the names of employees who work on more than one project.
c) Calculate the total salary expenditure for each department.
d) Display the names of employees along with the names of departments and their locations.
e) Find the names of employees who work on projects in the 'Marketing' department located in 'New York'.

Q9. Determine whether the given schedule   **10**   **5**   **2**

$$S = r1(X) \ w2(Y) \ w1(X) \ r3(Y) \ r2(X) \ r4(Z) \ w4(Z)$$

Satisfies the requirements of the Two-Phase Locking (2PL) protocol. Describe the lock acquisition and release sequence in this schedule with 2PL.

Q10. A) Draw the Serializable graphs for the schedule S1 and state whether each schedule is conflict   **10**   **5**   **3**
serializable or not. If a schedule is conflict serializable,
write down the order in which the transaction can be executed.
S1:r1(X);r2(Z);r1(Z);r3(X);r3(Y);w1(X);w3(Y);r2(Y);w2(Z);w2(Y)

B) Elaborate on recovery techniques (deferred and immediate) for single and multi-user transactions.

Q11. An ordered file of 1,00,000 records is stored in a disk with block size 2048 bytes. File records are   **10**   **6**   **3**
fixed and are un-spanned with record length 256 bytes. Suppose the index file has key field of 10
bytes and pointer of 6 bytes, Compare the performance of accessing data file directly and by
primary index.

Q12. Define dynamic hashing? Explain how dynamic hashing dynamically adjusts hash table size to   **10**   **6**   **2**
accommodate data growth.

End of Question Paper

VIT-AP UNIVERSITY

Final Assessment Test – Winter 2023-24 Semester - May 2024

Max Marks: **100**	Duration: **3 Hours**

Course Code: **CSE2007**	Course Title: **Database Management Systems**	
Question Paper ID: **009716**	Exam Type: **CLOSED BOOK**	School: **SCOPE**
Date: **11/05/2024**	Slot: **B1+TB1**	Session: **FN**

Keeping mobile phone/smart watch, even in 'off' position is treated as exam malpractice.

General Instructions if any:

1. "*fx* series" - non Programmable calculator is permitted: No
2. Reference tables permitted: No

Section - 1: Answer any 10 questions. (10 × 10 = 100 Marks)

		Marks	CO	BL

Q1. 1. You are tasked with designing a database for a library system. The library system has the following requirements: **10 · 2 · 2**

· The library contains multiple books, each identified by a unique ISBN, title, author(s), and publication year.
· Each book belongs to a specific category (e.g., Fiction, Non-fiction, Science, History).
· Library members can borrow books. Each member has a unique ID, name, address, and contact number.
· Books can be borrowed by multiple members, and each member can borrow multiple books.
· The library keeps track of the borrowing history, including the date a book was borrowed and the date it was returned.

Design an Entity-Relationship (ER) diagram for the library system based on the provided requirements. Include all necessary entities, attributes, relationships, and cardinalities.

Q2. Describe the components of the Three Schema Architecture and explain how they facilitate data independence in a database system. **10 · 1 · 2**

Q3. Convert the given ER diagram into Relational Mapping **10 · 2 · 4**
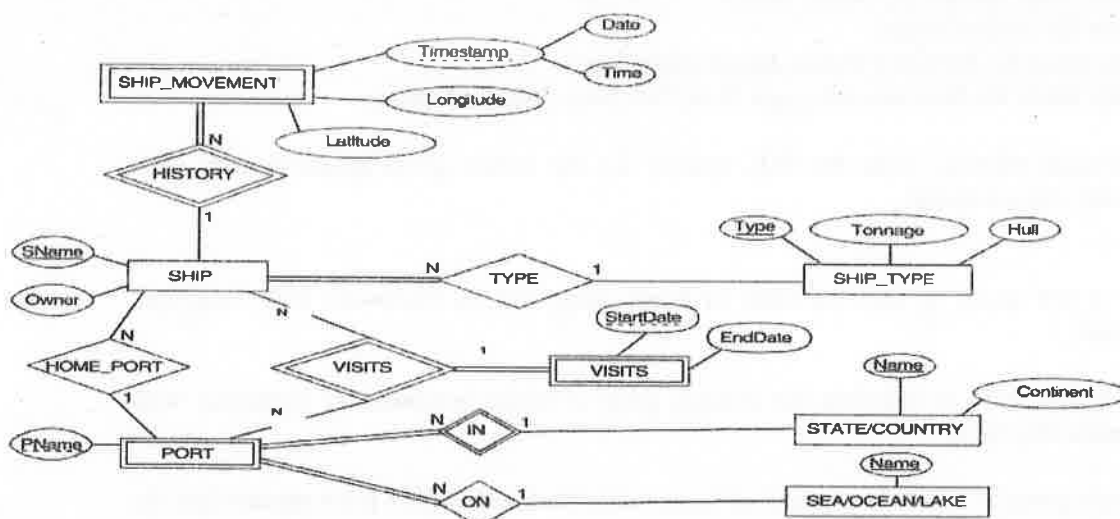
Q4. Consider the following relational schema representing a university database system: **10   2   2**

Student (student_id, name, major, GPA)
Course (course_id, title, department, credits)
Enrollment (student_id, course_id, semester, grade)

Perform the following queries using tuple relational calculus:

a) Retrieve the names of students who have not enrolled in any courses during the 'Spring 2024' semester.
b) Find the names of students who have never earned a grade below 'B' in any course they have taken.
c) Retrieve the titles of courses that have been taken by students majoring in 'Computer Science' with a GPA less than 3.0.
d) Find the names of students who have taken courses from at least two different departments.

Q5. a) Identify the highest normal form for the given relation below :   **10   3   3**

$R = (P, Q, R, S, T)$
$FD = (P \rightarrow QRST, QR \rightarrow PRT, S \rightarrow T)$

b) For the given relation and functional dependencies find all the candidate keys and prime attributes.

$R = (M, N, O, P)$
$FD = (M \rightarrow N, N \rightarrow O, O \rightarrow M)$

Q6. a) Find the Equivalence of functional dependencies sets given below   **10   3   3**
R(A,B,C,D)
$FD1 = (A \rightarrow B, B \rightarrow C, A \rightarrow C)$
$FD2 = (A \rightarrow B, B \rightarrow C, A \rightarrow D)$

b) Check whether the below given relation is in 3NF? If not explain the reason for not being in 3NF.
$R (G, H, I, J, K, L)$
$FD = (GH \rightarrow IJKL, HJ \rightarrow L)$

Q7. Consider the following relational schema representing a database for a library management system.   **10   4   2**

Books (Book ID, Title, Author ID, Genre, Price)
Authors (Author ID, Author Name)
Borrowers (Borrower ID, Borrower Name, Membership Type)
Loans (Loan ID, Book ID, Borrower ID, Loan Date, Due Date, Returned Date)

Using the provided schema, write the SQL queries for the below given questions with each involving at least one sub query:

a)   Write a sub query to find the titles of books borrowed by borrowers with 'Premium' membership type.

b)   Write a sub query to calculate the average price of books borrowed by borrowers with 'Premium' membership type.

c)   write a sub query to count the number of loans made for books with a price greater than the average price of all books.

c)   Write a sub query to list the names of authors who have written books with a price higher than the average price of all books.

Q8. Consider the following relational schema representing a database for a vehicle management system: **10**   **4**   **2**

Vehicles (VehicleID, Model, ManufacturerID, Year, Price)
Manufacturers (ManufacturerID, ManufacturerName, Country)
Orders (OrderID, CustomerID, VehicleID, OrderDate, Quantity)

a)   Write the SQL query and its corresponding relational algebra to draw a query tree for the following query:
How many vehicles of each model were ordered in 2023, along with their manufacturers and manufacturing years?

b)   Apply heuristic query optimization techniques to find the optimal query tree.

Q9. How does concurrency control in database systems address the challenges posed by issues such as   **10**   **5**   **2**
lost updates, temporary updates, incorrect summaries, and unrepeatable reads, ensuring data consistency and integrity in multi-user environments? Explain the issues with examples.

Q10. a) Check whether the given schedule S is conflict serializable or not. Using the precedence graph   **10**   **5**   **2**
find the execution order of different transaction.

T1	T2	T3	T4
			R(A)
	R(A)		
		R(A)	
W(B)			
	W(A)		
		R(B)	
	W(B)		

b) Differentiate between Deferred update and immediate update with suitable examples.

Q11. Consider that the data-file has 850000 records of employees. The employee-records are ordered   **10**   **6**   **4**
based on non-key field dept_no (i.e., department number) and size of that field is 6 bytes. Each of
the employee records has size of 120 bytes. Consider none of the records are allowed to cross block
boundaries. Consider clustering index structure for the data file has been created to speed up the
retrieval of the records where indexing field is based on the dept_no and a block pointer which is 6
bytes long. Assume the Block size is 10 Kilobytes. Consider, there are 50 employees working in
each of the department. How many block accesses will be needed to access a record from the data
file by searching the clustering index?

Q12.   **10**   **6**   **2**

Discuss the concept of internal hashing in a database management system (DBMS). Explain the
structure and functionality of internal hashing. Illustrate with examples how internal hashing
enhances the efficiency of data retrieval and storage management within a DBMS

End of Question Paper

Keeping mobile phone/smart watch, even in 'off' position is treated as exam malpractice.

General Instructions if any:

1. "*fx* series" - non Programmable calculator is permitted: No
2. Reference tables permitted: No

Note: Assume data wherever necessary. Any assumptions made should be clearly stated.

Section - 1: Answer any 10 questions. (10 × 10 = 100 Marks)

		Marks	CO	BL
Q1.	Consider the following case study and draw an ER diagram step by step. Represent participation type of each entity in the relation, constraints, and the cardinality of each relation.	10	2	4

An **online retail store** can use a relational database to manage product inventory, customer data, and orders. The 'Products' table would contain information about each product like id, name, price, number of items available in the store, and category id of the product. In each category, there are many products available. The 'Customers' table would store customer information, and the 'Orders' table maintain information about the customer, date of order, and total cost of order. The details of each order is maintained in 'OrderItem' table. This setup allows the store to maintain accurate records of customer transactions and inventory.

		Marks	CO	BL
Q2.	Explain DBMS component modules and their interactions with a suitable diagram.	10	1	1

Q3. Convert the following ER diagram into a relational data model included with constraints. 10 · 2 3
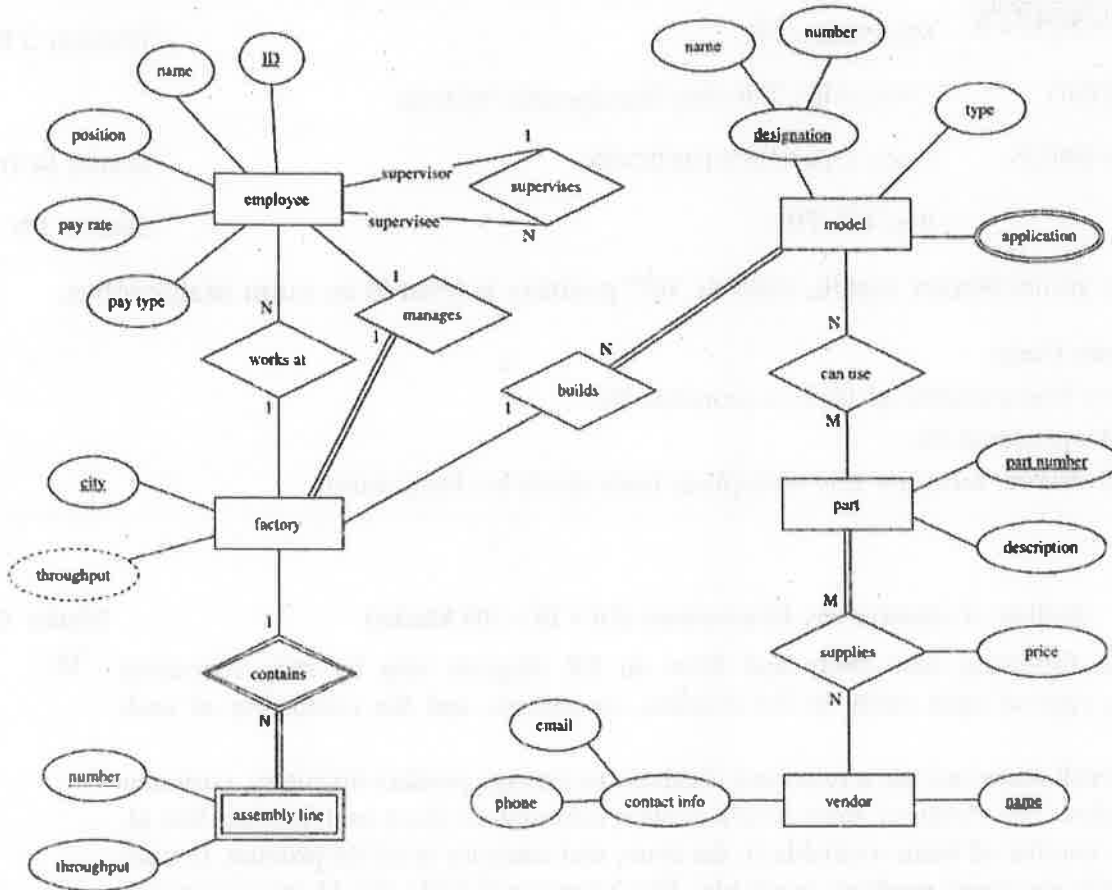
Q4. Consider the following GRADEBOOK relational schema describing the data for a grade 10 2 4
book of a particular instructor. The attributes A, B, C, and D of COURSES store grade cut-
offs.

CATALOG(Cno, Ctitle)
STUDENTS(Sid, Fname, Lname, Minit)
COURSES(Term, Sec_no, Cno, A, B, C, D)
ENROLLS(Sid, Term, Sec_no)

Specify the following queries in Relational Algebra on the GRADEBOOK database
schema.
a. Retrieve the names of students enrolled in the Automata class during the fall 2009 term
and Python class during the Winter 2010 using set operations.
b. Retrieve the Sid values of students who have enrolled in CSE2006 and CSE2007.
c. Retrieve the names of students who have not enrolled in any class.
d. Retrieve the names of students who have enrolled in all courses in the CATALOG table.

Q5. a. Consider the relation scheme R = {E, F, G, H, I, J, K, L, M, N} and the set of functional 10 3 3
dependencies
F = {EF -> G, F -> IJ, EH -> KL, K -> M, L -> N} on R. List all candidate keys of
R?

b. Find the canonical cover of the following set of functional dependencies. F= { A -> BC,
B -> C, A -> B,AB -> C }

Q6. a. Consider the relational schema R=(A,B,C,D,E) with its set of functional dependencies F = {A -> BC, C -> DE}. Decompose the relation R up to BCNF.   **10   3   3**

b. Find the highest normal form of a given relation R(A,B,C,D,E) with the following set of functional dependencies:
F = {A->D, B->A, BC->D, AC->BE}

Q7. Consider the following database schema: **10   4   3**

Depositer(customer_id,customer_name,customer_age)
Branch(branch_id,branch_name,branch_city)
Account(account_number,customer_id,branch_id)

a. Write an SQL Query and equivalent relational algebra expression to find all customer names who have account in either Vijayawada city or Hyderabad city.
b. Construct a Heuristic optimized query tree for the relational algebra expression found in (a). Delineate the process step by step.

Q8. Consider the following schema and answer the queries in SQL: **10   4   4**

Resort(resortNo, resortName, resortType, resortAddress, resortCity, numSuite)
Suite(suiteNo, resortNo, suitePrice)
Reservation(reservationNo, resortNo, visitorNo, checkIn, checkout, totalVisitor, suiteNo)
Visitor(visitorNo, firstName, lastName, visitorAddress)

a. List all details of all resorts in Mumbai which have number of suits greater than 30 and suite price between 1.5 Lakh to 2.0 Lakh.
b. List all first names of all visitors who have booked suites and their cost is above 3 Lakh.
c. Display all visitor's first names in the descending order of reservations whose reservations crossed certain threshold value given by user.
d. Identify all visitors who reserved all resorts of 'Mumbai' city.

Q9. Explain how serializabilty ensured by Two Phase Locking protocol? How the variants of Two Phase Locking protocol alleviates drawbacks of Basic 2PL with examples?   **10   5   1**

Q10. Explain Deferred update recovery protocol. From the following diagram, identify which   10   5   2
transactions are ignored, which transactions are redone and which (if any) are undone.
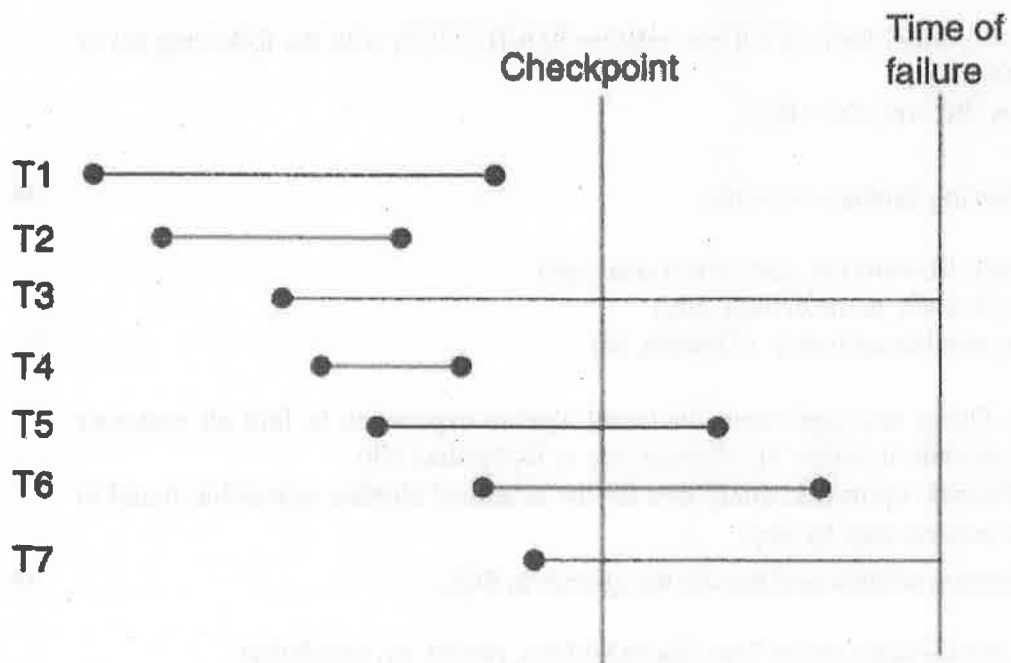
Q11. Define Indexing and What are the advantages of indexing in DBMS? How does multilevel   10   6   2
indexing improve the efficiency of searching an index file?

Q12. Explain Internal hashing technique and its' collision resolution approaches.   10   6   1

<div align="center">***End of Question Paper***</div>

Keeping mobile phone/smart watch, even in 'off' position is treated as exam malpractice.

General Instructions if any:

1. "*fx* series" - non Programmable calculator is permitted: No
2. Reference tables permitted: No

<div align="center">

Section - 1: Answer any 10 questions. (10 × 10 = 100 Marks)

</div>

		Marks	CO	BL

Q1. A) Name any five types of Database users and explain their interaction with the Database. — 10 | 1 | 1

B) Why Database is called self-describing? Explain the structure of the database catalog.

Q2. Draw the E-R Diagram for the below scenario (E-R diagram should have entity, attribute, and Relation, with cardinality). Write primary keys, foreign keys, and all other database constraints. — 10 | 1 | 4
Name, Release_year, Language, Age_group, Producer, Director and Duration can identify a movie. Actors can be identified by Name, Date of Birth, Address and work language. Movie_Theater can be identified by theatre Name, Established year, Address, Owner and ticket price. An actor will act in different movies with some characters and get remuneration. Movies will be released in some theatres and will be shown on multiple shows.

Q3. Consider the given Database schema for answering the queries using relational algebra. — 10 | 2 | 2
Relational schema:
Student(Sid, name, address, dob, gender, phone, email)
Course(Cid, C_type, name, credit, offered_by, syllabus_by)
Teacher(Tid, name, gender, phone, email, Dept)
Registration(Sid, Cid, semester, slot, venue)
Teach(Tid, Cid, semester, slot, venue)
Pre_Anty_Requisite(Cid, Pre_Cid, Anty_Cid)
Queries:
a) Find the students, who have registered less than 8 credits in the Programcore Course type in win23-24.
b) List the teacher who is teaching a course not offered by his department.
c) Find the course details, which are offered by the ECE department but the syllabus made by CSE faculty.
d) List the courses having prerequisites and anty-requisites.

Q4. a) How intersection can be implemented using other relational operations. Explain with an example. — 10 | 2 | 1
b) Differentiate between Natural join, Full Outer Join and Equi Join. Explain with examples.

Q5. A) Here is the FD set S={C→(B E); D →(A F); B →(A C); E →F}. Find all candidate keys using attribute closer and Find any 5 super keys. — 10 | 3 | 4
B) A relational schema R (A, B, C, D, E) have FD set P = {A → B, AB → C, D → AEC} and Q = {A → BC, D → AE}. Find out the FD equivalence between P and Q.

Q6. The relation X(P, Q, R, S) have FD set A={P→(Q, S); Q→R; P→R}; — 10 | 3 | 4
a) Find the minimal cover of A.
b) Find the current Normal form of X(before minimal cover), and normalise X up to 3NF.

Q7.	Consider the relational schema given below to answer the Queries using SQL. Employee(Eno, Name, Email, Office-location, Gender) Customer(Cno, name, email, address, phone, age) Product(Code, Name, Type, Size, Price, Vendor) Orders(OrderNo, Cno, Eno, Code, Type, Quantity, TotalPrice, OrderDate, Status) Payment(ReferenceNO, OrderNo, PayMethod, PaidAmount, DueAmount, PaymentDate) Queries: a) Find the customer names who made the order and it is handled by an employee from the same location. b) List the vendor-wise product name with type and price. c) List the customers, who made an order and paid the full amount. d) Find the product name with the code, which has been sold least quantity till date.	10	4	4
Q8.	Consider the relational schema given below and find the optimal query tree for the given query. Employee(Eno, Name, Email, Office-location, Gender) Customer(Cno, Name, email, address, phone, age) Product(Code, Name, Type, Size, Price, Vendor) Orders(OrderNo, Cno, Code, Type, Quantity, TotalPrice, OrderDate, Status) Payment(ReferenceNO, OrderNo, PayMethod, PaidAmount, DueAmount, PaymentDate) Query: SELECT C.Name FROM Customer C, Orders O, Payment P WHERE C.CNo = O.CNo and P.OrderNo= O.OrderNo and P.PaidAmount < P.DueAmount and O.OrderDate = P.PaymentDate and O.status= 'transite';	10	4	3

Q9. We have Transactions
$T_1(R(X); W(X); R(Y); W(Y))$,
$T_2(R(Z); W(Z); R(X))$ and
$T_3(R(Y); R(Z); R(X); W(Y); W(X))$.
If database system executes them as Schedule S= { $R_1(X)$, $R_2(Z)$, $W_2(Z)$, $W_1(X)$, $R_1(Y)$, $R_2(X)$, $R_3(Y)$, $R_3(Z)$, $R_3(X)$, $W_3(Y)$, $W_3(X)$, $W_1(Y)$}.

a)   List all the problems (Dirty read, lost update, unrepeatable read, incorrect summary) may occur in S. Explain.
b)   Find the Schedule S is conflict Serializable or not using a precedence graph.   **10 5 3**

Q10. Explain the shadow Paging database recovery system with a proper diagram.   **10 5 2**

Q11. Explain the Secondary Indexing method and Cluster Indexing method with a relevant diagram.   **10 6 1**

Q12. Relation Student has 4500 records with a record length 55 byte, block size is 2048 byte.   **10 6 3**
A) How many blocks are required to store the relation?
B) How many blocks are required to store the index table using primary indexing and how many block access are required to search any record? Where the index field size is 6 byte and the pointer size is 5 byte.

End of Question Paper

VIT-AP UNIVERSITY

Final Assessment Test – Winter 2023-24 Semester - May 2024	
Max Marks: **100**	Duration: **3 Hours**

Course Code: **CSE2007**	Course Title: **Database Management Systems**	
Question Paper ID: **009724**	Exam Type: **CLOSED BOOK**	School: **SCOPE**
Date: **15/05/2024**	Slot: **C2+TC2**	Session: **FN**

Keeping mobile phone/smart watch, even in 'off' position is treated as exam malpractice.

General Instructions if any:
1. "*fx* series" - non Programmable calculator is permitted: No
2. Reference tables permitted: No

Section - 1: Answer any 10 questions. (10 × 10 = 100 Marks)

		Marks	CO	BL
Q1.	Describe the main components of a database management system (DBMS). Explain the function of each component and how they interact with each other.	10	1	2
Q2.	Design an ER diagram for a hotel reservation system. Include entities such as guests(Guest_ID , Name, Address, Phone, Email), rooms (Room_Number, Type, Price, Capacity, Status), reservations(Room_Number, Check_in_date, Check_out_date, Total_Price), and payments(Amount, Payment_Date). Additionally, include a weak entity for room services ordered by guests. Specify attributes for each entity, and include details for the weak entity. find all relationships between the entities.	10	2	3
Q3.	Convert the following ER diagram into a relation schema	10	2	3
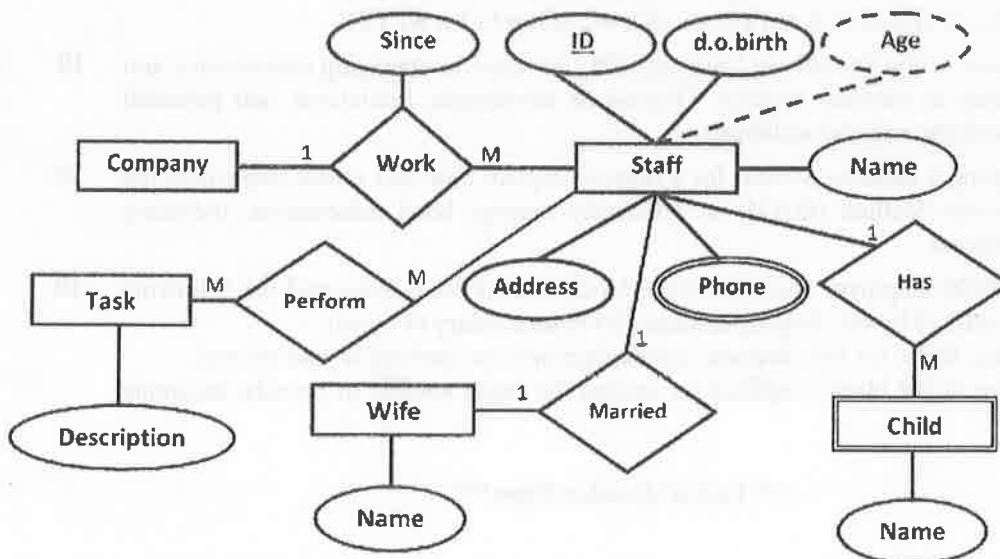
		Marks	CO	BL
Q4.	**Consider the following schema** **Student**(sid, sname, age) **Course**(cid, cname, credits) **Enrollment**(sid, cid, grade) a) Write a relational calculus expression to find the names of students who have enrolled in the course "Database Systems". b) Write a relational calculus expression to find the names of students who are older than 20 and have taken more than 4 courses. c) Find the names of students who have enrolled in a course worth more than 3 credits. d) Find the names of students who have taken all courses with a grade of 'A'.	10	2	4
Q5.	Consider a relation R(A, B, C, D, E) with functional dependencies {A → B, B → C, CD → E}. Determine if R is in BCNF and if not, decompose it into BCNF. Discuss the concept of dependency preservation.	10	3	3

Q6. Consider a relation R(A, B, C, D, E, F) with functional dependencies {A → B, B → C, C → D, D → E, E → F}.　　**10**　**3**　**1**
a) Determine if the set of functional dependencies is equivalent to {AB → C, BC → D, CD → E, DE → F}.
b) Find a minimal cover for the given set {A → B, B → C, C → D, D → E, E → F} for the above relation.

Q7. Consider the following SQL query to retrieve all applicants who intend to enroll in the Software　**10**　**3**　**1**
Engineering degree program, reside in Vijayawada, and attend a college with a rank of 25 or lower.
The Relations are as follows:
Applicants (Id, name, city, Cid)
College (Cid, Cname, rank)
Programme (Id, major)
SELECT A.NAME FROM APPLICANTS A, COLLEGE C, PROGRAMME P WHERE A.CID =
C.CID AND A.ID = P.ID AND A.CITY = 'Vellore' AND C.RANK < 25 AND P.MAJOR =
'Software Engineering';
Find the efficient execution plan for the above query and draw the optimized query tree.

Q8. Summarize the fundamentals of cost-based optimization in SQL query processing. Correlate the　**10**　**4**　**5**
cost-based optimization process, including how it works and the parameters taken into account
when producing query execution plans.

Q9. Consider the three transactions T1, T2, and T3, and the schedules S1 and S2 given below. Draw the　**10**　**5**　**6**
serializability (precedence) graphs for S1 and S2, and state whether each schedule is serializable or
not. If a schedule is serializable, write down the equivalent serial schedule(s).
T1: r1 (X); r1 (Z); w1 (X);
T2: r2 (Z); r2 (Y); w2 (Z); w2 (Y);
T3: r3 (X); r3 (Y); w3 (Y);
S1: r1 (X); r2 (Z); r1 (Z); r3 (X); r3 (Y); w1 (X); w3 (Y); r2 (Y); w2 (Z); w2 (Y);
S2: r1 (X); r2 (Z); r3 (X); r1 (Z); r2 (Y); r3 (Y); w1 (X); w2 (Z); w3 (Y); w2 (Y);

Q10. Evaluate the effectiveness of the Two-Phase Locking (2PL) protocol in managing concurrency and　**10**　**5**　**4**
ensuring data consistency in database systems. Discuss its advantages, limitations, and potential
trade-offs in terms of performance and scalability.

Q11. Suppose you are designing a database system for a library. Explain how you would implement the　**10**　**6**　**6**
Indexed Sequential Access Method (ISAM) to efficiently manage book information, including
indexing and record retrieval.

Q12. Given a database of 10000 employee records with a block size of 4096 bytes and the following　**10**　**6**　**4**
record format: EmployeeID (4 bytes), EmployeeName (50 bytes), Salary (8 bytes).
A) Calculate the blocking factor for this database and explain how you arrived at your answer.
B) Calculate the number of file blocks required for storing the entire number of records, assuming
unspanned organization.

End of Question Paper

Keeping mobile phone/smart watch, even in 'off' position is treated as exam malpractice

General Instructions if any:
1. "*fx* series" - non Programmable calculator are permitted : YES
2. Reference tables permitted : NO

PART – A: Answer any TEN Questions, Each Question Carries 10 Marks (10×10=100 Marks)

1	Answer the following. A) What is data abstraction? How this concept is implemented in DBMS. B) Write the responsibility of Database Administrator.	10M
2	Draw the E-R Diagram for the below scenario (E-R diagram should have entity, attribute, Relation with cardinality). Write primary keys, foreign keys, and all other database constraints. The flight database stores details about an airlines, flights, and seat bookings. Consider the following scenario list: • The airline has one or more airplanes. • An aeroplane has a model number, a unique registration number, and the capacity to take one or more passengers. • An aeroplane flight has a unique flight number, a departure airport, a destination airport, a departure date and time, and an arrival date and time. • Each flight is carried out by a single airplane. • A passenger has a name, surname, and a unique email address. • A passenger can book a seat on a flight.	10M
3	Consider the given Database schema for answering the queries using relational algebra. Relational schema: Movie (M_Name, Release_year, Language, Age_group, Producer, Director, Duration); Actors (A_Name, DOB, Gender, Address, Working_language); Movie_Theater (MT_Name, Established_year, Address, Owner, Ticket_price). Acted_on (A_Name, M_name, Remuneration, character); Release (M_name, MT_name, Number_of_show); Queries: a) Find all movie names released in 2020 in English or Hindi. b) List the theaters in Vijawarda showing the movie "Jurassic Park". c) Find the senior Most Female Actor. d) List the movie names, where remuneration is more than 10 lakh.	10M

4	Consider the Following ER diagram and map it into a relational model with proper constraints and justification.	10M
5	Given R(X, Y, Z, W) and Set of Functional Dependency FD = { X → Y, Y → Z, Z → X}. Find all candidate key(s) in above relation R using a given set of FDs.	10M
6	Let R (A, B, C, D, E, F, G, H) be a relational schema with FD set {C→A, B→D, CD→FG, G→FH, D→BE}. Find the current normal form and normalise R upto 3NF.	10M

7	Consider the relational schema given below to answer the Queries using SQL.
	Employee(Eno, Name, Email, Office-location, Gender)
	Customer(Cno, name, email, address, phone, age)
	Product(Code, Name, Type, Size, Price, Vendor)
	Orders(OrderNo, Cno, Eno, Code, Type, Quantity, TotalPrice, OrderDate, Status)
	Payment(ReferenceNO, OrderNo, PayMethod, PaidAmount, DueAmount, PaymentDate)
	Queries:
	  a) Find the average age of all customers.
	  b) List the vendor-wise product name with type and price.
	  c) List orders, where the paid amount is greater than the Due amount.
	  d) Find the product name with the code, which has been sold on 20th July 2024.
		10M

8	Consider the relational schema given in question 7, and find the optimal query tree for the given query. **Query:** SELECT E.Name, C.name, O.OrderNo. FROM Employee E, Customer C, Orders O, WHERE C.CNo = O.CNo and O.Eno = E.Eno and O.TotalPrice < 20k and E.Gender = 'male' and O.status= 'Pending';	10M
9	We have three Transactions and if the database system executes them as Schedule S= { $R_3(B)$, $R_3(C)$, $R_1(A)$, $W_1(A)$, $W_3(B)$, $W_3(C)$, $R_2(C)$, $R_1(B)$, $W_1(B)$, $R_2(B)$, $W_2(B)$, $R_2(A)$, $W_2(A)$}. Find this Schedule S is a conflict Serializable or not using a precedence graph.	10M
10	Explain the shadow Paging database recovery system with a proper diagram.	10M
11	Explain the Primary indexing method with an example and proper diagram.	10M

12	Relation Student has 3900 records with below schema.
	Student (RegNO number, Name varchar(20), Address varchar(40), Degree varchar(10), Department(5), CGPA number). If the block size is 2048 bytes.
	  A) Find the maximum size of each record possible.
	  B) Find the maximum possible size of the student table.
	  C) How many blocks are required to store the relation using an un-spaned record structure?
		10M

Keeping mobile phone/smart watch, even in 'off' position is treated as exam malpractice

General Instructions if any:
1. Numerical questions - non-Programmable calculator are permitted: YES /~~NO~~
2. Reference tables are permitted: ~~YES~~ / NO

PART – A: Answer any **TEN** Questions, Each Question Carries 10 Marks (10×10=100 Marks)

1. A. Explain the roles of three types of database users. **(5 M)**
   B. Differentiate between total participation and partial participation using examples. **(5 M)**

2. Consider the bank database given below and answer the following questions. **(10 M)**
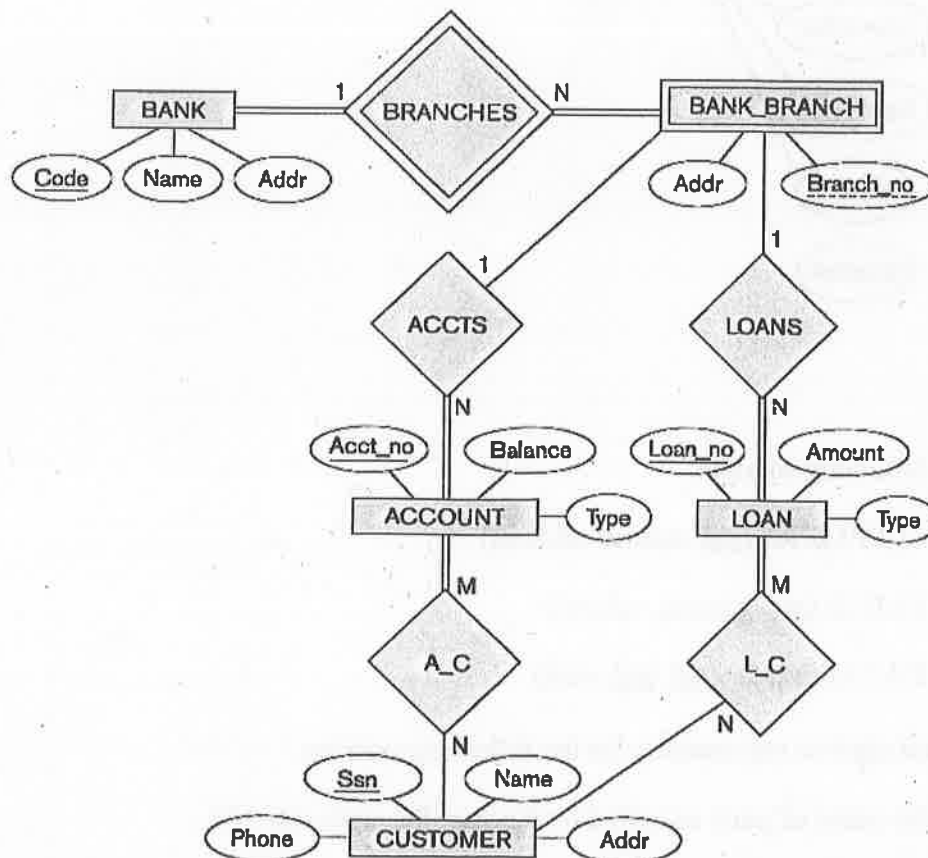
i. List the strong (regular) entity types in the ER diagram.
ii. Is there a weak entity type? If so, give its name, partial key, and identifying relationship.

iii. What relational model constraints do the partial key and the identifying relationship of the weak entity type impose in this diagram?

iv. Suppose that every customer must have at least one account but is restricted to at most two loans at a time, and that a bank branch cannot have more than 1,000 loans. How are these constraints represented using the (min, max) notation?

3. Using the standard procedure, convert the following ER diagram to relational schema. **(10 M)** Identify primary keys and foreign keys of the generated relations.
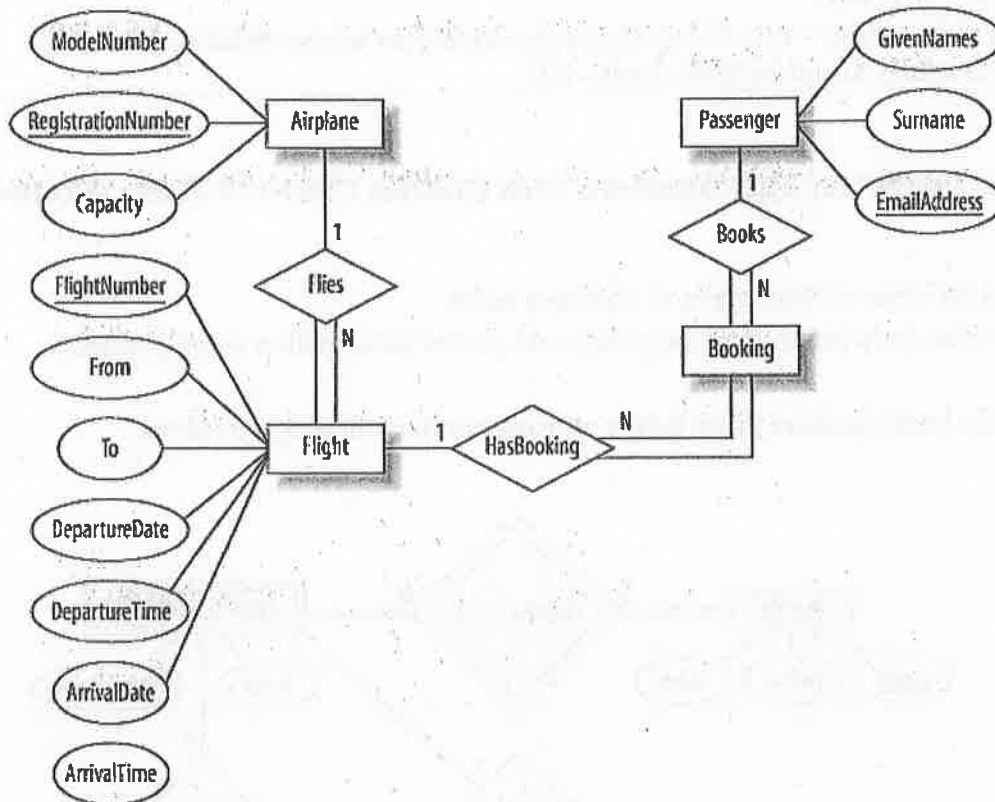
4. Consider the following schema: **(10M)**

        SUPPLIERS (sid, sname, address)

        PARTS (pid, pname, colour)

        CATALOGUE (sid, pid, cost)

Write relational algebra expressions for the following queries:

    i. Find the name of parts supplied by the supplier with sid=105.

    ii. Find the sids of suppliers who supply every part.

    iii. Find the names of suppliers supplying some green parts for less than Rs. 1000.

    iv. Find the sids of suppliers who supply some red parts and some green parts.

5. P and Q are two set of FDs for a relational schema R (A, B, C, D). P = {A → B, B → C, C → D} and Q = {A → BC, C → D}. Determine if P covers Q and if Q covers P. Additionally, check if P and Q are equivalent. **(10 M)**

6. Let R: (A, B, C, D, E), Rl: (A, D), R2: (A, B), R3: (B, E), R4: (C, D, E), and R5: (A, E). Let the FDs be: A → C, B → C, C → D, A → D, DE → C, CE → A. Check whether the decomposition of R into {Rl, R2, R3, R4, R5} is a lossless decomposition or not. **(10 M)**

7. Consider the schema for Company Database: **(10 M)**

> EMPLOYEE (SSN, Name, Address, Sex, Salary, SuperSSN, DNo)
> DEPARTMENT (DNo, DName, MgrSSN, MgrStartDate)
> DLOCATION (DNo, DLoc)
> PROJECT (PNo, PName, PLocation, DNo)
> WORKS_ON (SSN, PNo, Hours)

Write SQL queries to

   i.   Make a list of all project numbers for projects that involve an employee whose last name is 'Scott', either as a worker or as a manager of the department that controls the project.
   ii. Show the resulting salaries if every employee working on the 'IoT' project is given a 10 percent raise.
   iii. Find the sum of the salaries of all employees of the 'Accounts' department, as well as the maximum salary, the minimum salary, and the average salary in this department
   iv. Retrieve the name of each employee who works on all the projects controlled by department number 5.
   v.   For each department that has more than five employees, retrieve the department number and the number of its employees who are making more than Rs. 6,00,000.

8. Consider three tables **(10 M)**

> COURSE (CNO, CNAME, CREDITS)
> STUDENT (ROLLNO, NAME, ADDRESS, SEM) and
> ENROLLMENT (CNO, ROLLNO, GRADE). Foreign keys have the same name as primary keys. Identify one initial canonical query tree for the following SQL expression and show the steps to optimize it using heuristics.

> SELECT S.NAME, S.ADDRESS, E.GRADE
> FROM COURSE C, STUDENT S, ENROLLMENT E
> WHERE S.ROLLNO =E.ROLLNO AND C.CNO = E.CNO AND CNAME='PDBD'

9. A. Illustrate lost-update and dirty-read problems with suitable examples. **(5 M)**
    B. Describe the different states of a transaction with the help of a neat diagram. **(5 M)**

10. Consider the following schedules. Determine whether each schedule is strict, cascadeless, **(10 M)** recoverable, or nonrecoverable. Justify your assumptions.

$S1$: $r_1(X)$; $r_2(Z)$; $r_1(Z)$; $r_3(X)$; $r_3(Y)$; $w_1(X)$; c_1; $w_3(Y)$; c_3; $r_2(Y)$; $w_2(Z)$; $w_2(Y)$; c_2;

$S2$: $r_1(X)$; $r_2(Z)$; $r_1(Z)$; $r_3(X)$; $r_3(Y)$; $w_1(X)$; $w_3(Y)$; $r_2(Y)$; $w_2(Z)$; $w_2(Y)$; c_1; c_2; c_3;

11. Explain the different types of single-level ordered indices with suitable examples. **(10 M)**

12. Consider an EMPLOYEE file with 10000 records where each record is of size 80 bytes. **(10 M)** The file is sorted on employee number (15 bytes long), which is the primary key. Assuming unspanned organization, block size of 512 bytes and block pointer size of 5 bytes, compute the number of block accesses needed for retrieving an employee record based on employee number using binary search if
   i.   No index is used
   ii.   Primary-level index is used

QP MAPPING

Q. No.	E/A/T	Module Number	Marks	BL	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped
Q1	E	M1	10	2	CO1	1, 12	1, 4	1
Q2	T	M1	10	3	CO1	1, 2, 3, 12	1, 4	1
Q3	T	M2	10	3	CO2	1, 2, 3, 12	1, 4	1
Q4	A	M2	10	3	CO2	1, 2, 3, 12	1, 4	1
Q5	A	M3	10	3	CO3	1, 2, 3, 12	1, 4	1
Q6	A	M3	10	3	CO3	1, 2, 3, 12	1, 4	1
Q7	T	M4	10	3	CO4	1, 2, 3, 12	1, 4	1
Q8	A	M4	10	3	CO4	1, 2, 3, 12	1, 4	1
Q9	E	M5	10	2	CO5	1, 12	1, 4	1
Q10	A	M5	10	3	CO5	1, 2, 3, 12	1, 4	1
Q11	E	M6	10	2	CO6	1, 12	1, 4	1
Q12	A	M6	10	3	CO6	1, 2, 3, 12	1, 4	1

VIT-AP UNIVERSITY	Regular Arrear Examinations (2023-24) - July 2024	
	Maximum Marks: 100	Duration: 3 Hours
Course Code: CSE2007	Course Title: **Database Management Systems**	
Set No: 1	Exam Type : **Close Book**	School: SCOPE
Date: 09/08/2024	Slot: F	Session: FN

Keeping mobile phone/smart watch, even in 'off' position is treated as exam malpractice

General Instructions if any:

1. *"fx series"* - non Programmable calculator are permitted : NO
2. Reference tables permitted : NO

PART – A: Answer any TEN Questions, Each Question Carries 10 Marks (10×10=100 Marks)

1. Draw an ER diagram for the University database following the below schema structure using with all different components like strong and weak entities, key attributes, relationships. **(10 M)**
   College (CName, COffice, CPhone)
   Dept (DName, DNo., DPhone)
   Course (CId, CName, Credit, CLevel)
   Instructor (IId, IName, IOffice, IPhone, Age, DOB)
   Student (Regd no., Name, DNo., Phone, Age, DOB)
   Section (Sid, SName, Sem, Slot)

2. Suppose you're assigned to design a database system for an e-commerce organization. How would you categorize and elucidate the duties and functions of diverse database users within the organization? Explain how the three-schema architecture facilitates the e-commerce company in modifying their database structure to accommodate new data needs without disrupting their current applications. **(10 M)**

3. Suppliers (sid, sname, address) **(10 M)**
   Parts (pid, pname, color)
   Catalog (sid, pid, cost)
   Write the following queries in relational algebra
   a) Find the pids of parts supplied by at least three different suppliers.
   b) Find the pids of the most expensive parts supplied by suppliers named Sunil

4. Consider the following schema **(10 M)**
   Student (sid, sname, age)
   Course (cid, cname, credits)
   Enrollment (sid, cid, grade)
   a) Write a relational calculus expression to find the names of students who have enrolled in the course "Database Systems".
   b) Write a relational calculus expression to find the names of students who are older than 20 and have taken more than 4 courses.

5. Consider a relationship R= (P, Q, R, S, T, X) on which the following FDs hold {P→Q, QR→S, T→R, S→P}. Find all the possible candidate key and super key attributes for the relation R. **(10 M)**

6. For the relation R(A, B, C, D), the decomposed relations are D={AB, BC, CD} and the FDs hold: {A→B, B→C, C→D, D→A} present in the relation R. Check that the given decomposed relations are dependency preservation or not ? **(10 M)**

7. Consider the following relations **(10 M)**
   Employee (Fname, Lname, SSN, Salary, Bdate, Address)
   Project (Pname, Pnumber, Plocation, Dnum)
   Works (ESSN, Pno, Hours)

   Write the SQL query and it's corresponding Relation algebra to draw an initial query tree for the query Find the last name of employees born after 1957 who worked on a project named "Women Empowerment". Then apply heuristic query optimization on the above initial query tree to find the optimal query tree.

8. Consider the following relational schema. **(10 M)**
   Employee (empno,name,office,age)
   Books (isbn,title,authors,publisher)
   Loan (empno, isbn,date)
   Write the appropriate SQL queries for the following questions.
   a) Find the names of employees who have borrowed more than five different books published by McGraw-Hill?
   b) Find the names of employees who have borrowed all books Published by McGraw-Hill?

9. How does concurrency control in database systems address the challenges posed by issues such as lost updates, temporary updates, incorrect summaries, and unrepeatable reads, ensuring data consistency and integrity in multi-user environments? Explain the issues with examples. **(10 M)**

10. Consider the three transactions T1, T2, and T3, and the schedules S1 and S2 given below. Draw the serializability (precedence) graphs for S1 and S2, and state whether each schedule is serializable or not. If a schedule is serializable, write down the equivalent serial schedule(s). **(10 M)**
    T1: r1 (X); r1 (Z); w1 (X);
    T2: r2 (Z); r2 (Y); w2 (Z); w2 (Y);
    T3: r3 (X); r3 (Y); w3 (Y);
    S1: r1 (X); r2 (Z); r1 (Z); r3 (X); r3 (Y); w1 (X); w3 (Y); r2 (Y); w2 (Z); w2 (Y);
    S2: r1 (X); r2 (Z); r3 (X); r1 (Z); r2 (Y); r3 (Y); w1 (X); w2 (Z); w3 (Y); w2 (Y);

11. Discuss the concept of internal hashing in a database management system (DBMS). Explain the structure and functionality of internal hashing. Illustrate with examples how internal hashing enhances the efficiency of data retrieval and storage management within a DBMS. **(10 M)**

12. Consider a database of fixed-length records, stored as an ordered file. The database has records, with each record being bytes, of which the primary key occupies bytes. The data file is block-aligned in that each data record is fully contained within a block. The database is indexed by a primary index file, which is also stored as a block-aligned ordered file. The figure below depicts this indexing scheme. **(10 M)**

QP MAPPING

Q. No.	E/A/T	Module Number	Marks	BL	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped
Q1	E	1	10	1	1	1	1	1
Q2	E	1	10	2	1	1	1	1
Q3	A	2	10	2	2	1,2	2	1
Q4	A	2	10	2	2	1,2	2	1
Q5	A	3	10	3	3	1,2,3,5	1,2	1
Q6	A	3	10	3	3	1,2,3,5	1,2	1
Q7	T	4	10	3	4	1,2,3,5	1,2	1
Q8	A	4	10	3	4	1,2,3,5	1,2	1
Q9	T	5	10	3	5	1,5	1,2	1
Q10	A	5	10	2	5	1,5	1,2	1
Q11	A	6	10	2	6	1,3,5	1,2	1
Q12	T	6	10	4	6	1,3,5	1,2	1

VIT-AP
UNIVERSITY
Apply Knowledge. Improve Life!

QUESTION PAPER

Name of the Examination: Short Summer-1-2023 – FAT

Course Code: **CSE2007**

Course Title: **Database Management system**

Date of Exam: 24/06/2023 (FN) (D1)

Duration: **120 Min**

Total Marks: 60 Marks

Instructions:

1. Answer all the questions.

6x10=60M

1. Write the SQL Queries for the following using the given database?
   Primary Keys: Salesman_id in Salesman table
   Customer_id in Customers table

Salesman

salesman_id	name	city	commission
5001	James Hoog	New York	0.15
5002	Nail Knite	Paris	0.13
5005	Pit Alex	London	0.11
5006	Mc Lyon	Paris	0.14
5003	Lauson Hense		0.12
5007	Paul Adam	Rome	0.13

Customer

customer_id	cust_name	city	grade	salesman_id
3002	Nick Rimando	New York	100	5001
3005	Graham Zusi	California	200	5002
3004	Fabian Johnson	Paris	300	5006
3007	Brad Davis	New York	200	5001
3009	Geoff Cameron	Berlin	100	5003
3008	Julian Green	London	300	5002
3003	Jozy Altidore	Moscow	200	5007
3001	Brad Guzan	London		5005

   I.   Find those salesmen with all information who gets the commission within a range of 0.12 and 0.14. **(2.5M)**

   II. Find those salesmen with all information whose name containing the 1st character is 'N' and the 4th character is 'l' and rests may be any character. **(2.5M)**

   III. Find the highest grade for each of the cities of the customers. **(2.5M)**

   IV. Find the name and city of those customers and salesmen who lives in the same city. **(2.5M)**

2. Convert the following ER diagram to Relational Schema: **(10M)**

3.
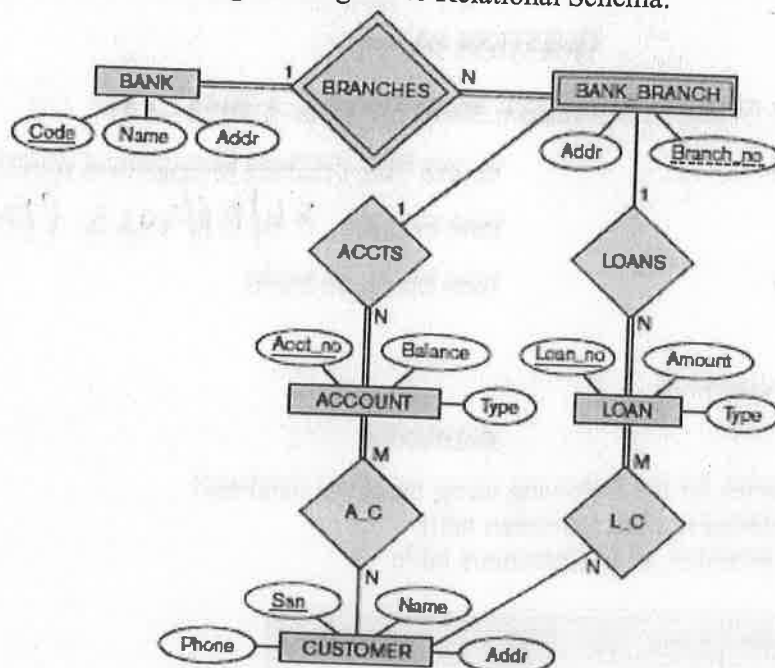
a. Consider the relation scheme R = {E, F, G, H, I, J, K, L, M, M} and the set of functional dependencies EF -> G, F -> IJ, EH -> KL, K -> M, L -> N on R. Find the candidate key for R?
(5M)

b. Find the canonical cover of the following set of functional dependencies. For the given relation R= {A, B, C, D, E, I}
F= {A → C, AB → C, C → DI, CD → I}
(5M)

4.

a. Discuss in detail two phase locking protocol mechanism?
(5M)

b. A relation R {C,S,J,D,P,Q,V} with following set of FDs:

    C->CSJDPQV, SD->P, JP-> C, J->S

This is decomposed as R1(C,J,D,Q,V), R2(J,S), R3(S,D,P). Find whether it is a lossless or lossy decomposition.
(5M)

5.

a. Justify why shadow paging technique does not use Undo/Redo technique? **(5M)**
b. Differentiate Primary Index and Secondary index with appropriate example? **(5M)**

6.

a. Describe the Process for heuristics optimization.
b. Given this database schema: **(5M)**
   •Product (pid, name, price)
   •Purchase (pid, cid, store)
   •Customer (cid, name, city)
(5M)

Draw the query tree for the below SQL query using the heuristic process.

SELECT DISTINCT x.store FROM Purchase x, Customer y WHERE x.cid = y.cid and y.city = 'Seattle';

QUESTION PAPER

Name of the Examination: FAT (WINTER 2022-2023)

Course Code: CSE2007 **Course Title: Database Management Systems**

Slot: I **Date of Exam:** 18/05/2023 (AN) (P2)

Duration: 120 Mins **Total Marks: 60**

1. Given the schema R(A,B,C) and S(D,E,F,G) and the following query SELECT E,B FROM S,R WHERE R.C=S.G AND R.A>5 AND S.E<3. Write the optimized relational algebra and draw the tree using heuristics in query optimization. **(10 Marks)**

2. Given the following schedules S1: $R_1(X)$ $R_1(Y)$ $R_2(X)$ $R_2(Y)$ $W_2(Y)$ $W_1(X)$ and S2: $R_1(X)$ $R_2(X)$ $R_2(Y)$ $W_2(Y)$ $R_1(Y)$ $W_1(X)$. Where Ri(A) means i^{th} transaction reads data A and Wi(A) means i^{th} transaction writes data A. Check whether the above schedules are conflict serializable or not.

   (10 Marks)

3. Given the following schedule S: $R_1(X)$ $W_2(X)$ $R_3(X)$ $W_1(X)$ $W_3(X)$. Where Ri(A) means i^{th} transaction reads data A and Wi(A) means i^{th} transaction writes data A. Check whether the schedule is view serializable or not. If yes then how many serial schedules are possible and what are they. **(10 Marks)**

4.
   a. Consider the following log consisting transactions T1, T2, and T3: **(5 Marks)**
      i. (Start, T1);
      ii. (Write, T1, P, 400, 600);
      iii. (Write, T1, Q, 400, 400);
      iv. (Commit, T1);
      v. (Start, T2);
      vi. (Write, T2, P, 600, 550);
      vii. (Write, T2, Q, 400, 450);
      viii. (Start, T3);
      ix. (Write, T3, P, 450, 600);
      x. (Write, T3, Q, 450, 400);
      xi. (Commit, T3);

   If the schedule crashes just after step X, then what will be the order of actions performed during the recovery process for undo logging and redo logging.

b. Suppose, we have to create an indexing on a column of a table, that is sorted but contains duplicate values. What kind of indexing we can do. With example explain the merits and demerits of this indexing technique. Explain how many of block access are required to search a data from the table using this indexing technique. **(5 Marks)**

5. Given a table with 50000 records, and 20 columns of each size 8 byte except the primary key which is of 10 byte. The block size is 2MB. Both primary indexing and secondary indexing is done on the table. The block pointer requires 6 bytes of space. Records are stored using unspanned mapping. The secondary indexing is done on a column which also contains only unique values in the table. Calculate the number of block access required to retrieve a record based on a search key using. **(10 Marks)**
    a. primary indexing
    b. secondary indexing.

6. Given a table with 20200 records, where the size of each record is 160 bytes and the size of each column is 16byte. The block size is 1.2 MB. Records are stored in spanned manner. Indexing done up to 3 levels on a unsorted column of the table. Calculate the number of block access required for searching a record using the indexing. The indexing is done on the column which contains only unique values. The block pointer size is one fourth of the size of the search key. **(10 Marks)**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	CO3	PO1, PO2, PO3, PO4			10
Q2	5	CO5	PO1, PO5			10
Q3	5	CO5	PO1, PO5			10
Q4	5	CO5	PO1, PO5			5+5 =10
Q5	6	CO6	PO1, PO3, PO5			10
Q6	6	CO6	PO1, PO3, PO5			10

Course Code: CSE 2007   **Course Title: Database Management Systems**

Set number: 2   Date of Exam: 13/05/2023 (AN)

Duration:120 min   Total Marks: 60   (G₁₂)

Instructions:

1. Assume data wherever necessary.

2. Any assumptions made should be clearly stated.

Q1. SELECT E. Lname

FROM EMPLOYEE E, WORKS_ON W, PROJECT P

WHERE P. Pname='Aquarius' AND P. Pnumber= W. Pno AND E. Essn=W. ssn

AND E. Bdate > '1957-12-31'

For the following query, prepare the initial (canonical) query tree, then show how the query tree is optimized by the use of heuristic query optimization.   **10M**

Q2. a. Check whether the given schedule S is conflict serializable or not-

$S : R_1(A) , R_2(A) , R_1(B) , R_2(B) , R_3(B) , W_1(A) , W_2(B)$ **5M**

b. Check whether the given schedule S is view serializable or not.   **5M**

T1	T2	T3	T4
R (A)			
	R (A)		
		R (A)	
			R (A)
W (B)			
	W (B)		
		W (B)	
			W (B)

Q3 a. Define database recovery. Explain the differences between Deferred update and Immediate update modifications in database recovery.   **7M**

  b. Explain 2PL protocol.   **3M**

Q4 a. Explain properties of transactions with examples.   **5M**

  b. Explain the advantages of shadow paging.   **5 M**

Q5 a. Define Indexing and illustrate the importance of Secondary Index.   **5M**

  b. Explain sequential file organization and the methods used to store file in hard disk.   **5M**

Q6. a. Table T has 7000 records with record length 40byte, block size is 1048byte. There are 2000 distinct values for indexing field.

How many blocks are required to store the table using simple cluster indexing? How many blocks are required to store the index table? Where index field size is 7byte and pointer size is 6byt.   **5M**

b. Explain Clustered indexing technique with diagram.   **5M**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	4	4	1,2,3,5			10
Q2	5	5	1,5			10
Q3	5	5	1,5			10
Q4	5	5	1,5			10
Q5	6	6	1,3,5			10
Q6	6	6	1,3,5			10

Course Code: CSE2007
Set number: 3
Duration: 120 Minutes

Course Title: Database Management Systems
Date of Exam: 15/05/2023 (FN) (A1)
Total Marks: 60 Marks

Instructions:
1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

---

1) Consider the transactions T1, T2, and T3 and the schedules S1 and S2 given below. **[10 M]**

T1: r1(X); r1(Z); w1(X); w1(Z)

T2: r2(Y); r2(Z); w2(Z)

T3: r3(Y); r3(X); w3(Y)

S1: r1(X); r3(Y); r3(X); r2(Y); r2(Z); w3(Y); w2(Z); r1(Z); w1(X); w1(Z)

S2: r1(X); r3(Y); r2(Y); r3(X); r1(Z); r2(Z); w3(Y); w1(X); w2(Z); w1(Z)

Check whether the Schedules S1 or S2 are conflict equivalent or not?
Give Justification to your answer with neat explanation.

2) Give an example of a schedule with a deadlock? Describe the techniques to prevent from deadlocks? **[10 M]**

3) Consider the following schedule, where each transaction is assumed to commit. **[10 M]**
S: r1(y) r3(y) r1(x) w2(x) r2(y) w3(x) w2(y)

Apply the rigorous two-phase locking protocol to the schedule. Is the schedule accepted?

4) a) When is it preferable to use a dense index rather than a sparse index? Explain your answer. **[5 M]**

b) Is it possible in general to have two primary indices on the same relation for different search keys? Explain your answer **[5 M]**

5) The following Figure shows the log corresponding to a particular schedule at the point of a system crash for four transactions T1, T2, T3, and T4. Suppose that we use the immediate update protocol with checkpointing. Describe the recovery process from the system crash.

Specify which transactions are rolled back, which operations in the log are redone and which (if any) are undone, and whether any cascading rollback takes place. **[10 M]**

[start_transaction, T_1]
[read_item, T_1, A]
[read_item, T_1, D]
[write_item, T_1, D, 20, 25]
[commit, T_1]
[checkpoint]
[start_transaction, T_2]
[read_item, T_2, B]
[write_item, T_2, B, 12, 18]
[start_transaction, T_4]
[read_item, T_4, D]
[write_item, T_4, D, 25, 15]
[start_transaction, T_3]
[write_item, T_3, C, 30, 40]
[read_item, T_4, A]
[write_item, T_4, A, 30, 20]
[commit, T_4]
[read_item, T_2, D]
[write_item, T_2, D, 15, 25] ◄——— System crash

6) For the relation R(ABCDEFGH) with FD's= {CH->G, A->BC, B->CHF, E->A, F->EG such that F+ is exactly the set of FDs that hold for R.} Consider the FDs given in above question. The relation R is in which Normal form and Explain why? **[10 M]**.

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	4	1	1	-	10 M
Q2	5	5	1, 5	1, 5	-	10 M
Q3	5	5	1, 5	1, 5	-	10 M
Q4	5	5	1, 5	1, 5	-	10 M
Q5	6	6	1, 3, 5	1, 3, 5	-	10 M
Q6	6	6	1, 3, 5	1, 3, 5	-	10 M

VIT-AP
UNIVERSITY
Apply Knowledge. Improve Life!

QUESTION PAPER

Name of the Examination: WINTER 2022-2023 – FAT

Course Code: CSE2007
Set number: 4
Duration: 120 Mins

Course Title: Database Management Systems
Date of Exam: 16/05/2023 (FN) (B₁)
Total Marks: 60

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. A car sales company offers its customers the option of purchasing a car using repayment schemes of over more than 5 years. A customer may purchase one or more cars under the scheme. When a customer purchases a car, he or she gets to choose one of the repayment scheme options for each purchase or may pay the whole amount in one go, meaning a repayment scheme option may be selected by one or more customers or may not be selected at all.

- A customer will have a unique customer number, name, address, and phone number. A repayment scheme option will have a unique repayment scheme number, name and the number of years for repayment.
- A car will have a unique car code, make, model, year of manufacture and price. The company pays commission on sales of cars to its salespeople and therefore it will also keep information about each sale which will include a unique sales ID and commission to be paid on that sale.
- The company will also want to know who made the sales and the car involved in that transaction so that the commission can be determined.
- A salesperson will have a unique salesperson ID, name, phone number and email address.

Analyze the above scenario and answer the following,

a) Which is the best database architecture model would you suggest to maintain these data? Why?
b) Design an Entity Relationship Diagram. **(10 M)**

Q2. Consider an online UPI payment as a test case. Draw the state transition diagram of different transaction execution and apply the appropriate steps that happen in an online UPI payment and explain. **(10M)**

Q3. Explain the following problems in concurrency control.

a) The Dirty read problem
b) The incorrect summary problem
c) Unrepeatable read problem

Suggest corrective measures to solve the problems **(10M)**

Q4. Describe a two-phase locking technique for concurrency control. Explain. How does it guarantee serializability? Consider 2 transactions T1 and T2 with process A, B, C, draw a strict 2 phase locking protocol structure mentioning the growing and shrinking phase.

(10 M)

Q5. Explain with a suitable example how the issues in transactions are solved using ACID properties.

(10M)

Q6. Explain the different types of file organization. Suggest the best file organization for huge files.

(10 M)

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	1	3	1,2,3,4	1,4	1	10
Q2	5	5	1,5	1,4	1	10
Q3	5	5	1,5	1,4	1	10
Q4	5	5	1,5	1,4	1	10
Q5	6	6	1,3,5	1,4	1	10
Q6	6	6	1,3,5	1,4	1	10

QUESTION PAPER

Name of the Examination: WINTER 2022-2023 – FAT

Course Code: CSE2007

Course Title: Database Management system

Set number: SET- 5

Date of Exam: 19/05/2023 (FN)(CE₁)

Duration: 120 Min

Total Marks: 60 Marks

Instructions:

1. Answer all the questions.
2. Any assumptions made should be clearly stated.
3. Read the questions carefully

6x10=60M

Q1. How can you explain the concept of Integrity Constraints in a relational database management system, and what are some examples of how they ensure data accuracy and consistency?   **10M**

Q2. **a.** Illustrate the benefit strict two-phase locking provide? List the disadvantages?   **5M**

b. Draw a state diagram and discuss the typical states that a transaction goes through during execution.   **5M**

Q3. Check whether the given schedule S is conflict serializable and recoverable or not?

T1	T2	T3	T4
	R(X)		
		W(X)	
		Commit	
W(X)			
Commit			
	W(Y)		
	R(Z)		
	Commit		
			R(X)
			R(Y)
			Commit

10M

Q4. Describe Undo and Redo operations of recovery technique with an example. Consider a simple checkpoint protocol and the following set of operations in the log with immediate update using undo/redo logging. Explain how the system recovers from the crash.   **10M**

```
[start_transaction, T₁]
[read_item, T₁, A]
[read_item, T₁, D]
[write_item, T₁, D, 20, 25]
[commit, T₁]
[checkpoint]
[start_transaction, T₂]
[read_item, T₂, B]
[write_item, T₂, B, 12, 18]
[start_transaction, T₄]
[read_item, T₄, D]
[write_item, T₄, D, 25, 15]
[start_transaction, T₃]
[write_item, T₃, C, 30, 40]
[read_item, T₄, A]
[write_item, T₄, A, 30, 20]
[commit, T₄]
[read_item, T₂, D]
[write_item, T₂, D, 15, 25]   ◄──── System crash
```

Q5. Consider a disk with block size B = 512 bytes. A block pointer is P = 6 bytes long. A file has r = 30,000 EMPLOYEE records of fixed length. Each record has the following fields: Name (30 bytes), Ssn (9 bytes), Department_code (9 bytes), Address (40 bytes), Phone (10 bytes), Birth_date (8 bytes), Sex (1 byte), Job_code (4 bytes), and Salary (4 bytes, real number). An additional byte is used as a deletion marker.   **10M**

    a. Calculate the record size R in bytes.

    b. Calculate the blocking factor bfr and the number of file blocks b, assuming an unspanned organization.

    c. Calculate the number of block accesses required to perform binary search with SSN as primary key using primary index

Q6.   **a.** Why can we have at most one primary index on a file, but several secondary indexes?   **5M**

    b. Illustrate an example describes dense index, and nondense (sparse) index?   **5M**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	1	CO1	PO1, PO4			10
Q2	5	CO5	PO1, PO5			10
Q3	5	CO5	PO1, PO5			10
Q4	5	CO5	PO1, PO5			10
Q5	6	CO6	PO1, PO3, PO5			10
Q6	6	CO6	PO1, PO3, PO5			10

QUESTION PAPER

Name of the Examination: WINTER 2022-2023 – FAT

Course Code: CSE2007

Course Title: Database Management Systems

Set number: 6

Date of Exam: 18/05/2023 (FN) (D1)

Duration: 120 Minutes

Total Marks: 60

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. **i.** Let R (A B C D E) be a relation with the following set of functional dependencies:
F: {CE -> D, D -> B, C -> A}
In this regard, determine the highest normal form satisfied by R. Justify your answer. **(5M)**

ii. Let R (A B C D E) be a relation with the following set of functional dependencies:
F: {BC -> ADE, D -> B}
In this regard, determine the number of super keys that can be formed. **(5M)**

Q2. Consider the below given serial schedule 'S' that involves three transactions T1, T2, and T3 respectively.

T1	T2	T3
READ(A)		
WRITE(B)		
	READ(A)	
	WRITE(B)	
		READ(A)
		WRITE(B)

In this regard, determine the number of schedules that are conflict equivalent to 'S'. In addition, determine the number of schedules that are view equivalent to 'S'. **(10M)**

Q3. Consider the following two transactions T1 and T2 respectively:
T1: READ(X) READ(Y) WRITE(X) WRITE(Y) COMMIT
T2: READ(X) WRITE(X) COMMIT
In this regard, determine the number of cascadeless schedules that can be formed using T1 and T2. **(10M)**

Q4. i. 'If a schedule is not conflict serializable, then Two-Phase Locking protocol can't be applied on it'. Validate or invalidate the statement using an example. **(5M)**

ii. Why is shadow paging called as a No undo/ No redo recovery mechanism? Explain. **(5M)**

Q5. Consider an unordered file of 1,00,000 records with a record size of 100 bytes stored on disk blocks of 4096 bytes using the unspanned record organization technique. How many blocks would be needed to store this file? On an average, how many blocks would have to be accessed to retrieve a particular record? **(10M)**

Q6. Consider an ordered file with the following specifications:

- Number of records= 10^8
- Record size= 400 bytes
- Spanned record organization technique used for allocating disk space to the file
- Key field size= 16 bytes
- Disk block pointer size= 4 bytes
- Disk block size= 4096 bytes

Then, how many blocks are needed to store the data file and index file respectively? **(10M)**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	4	1	1	-	10
Q2	5	5	1, 5	1, 5	-	10
Q3	5	5	1, 5	1, 5	-	10
Q4	5	5	1, 5	1, 5	-	10
Q5	6	6	1, 3, 5	1, 3, 5	-	10
Q6	6	6	1, 3, 5	1, 3, 5	-	10

QUESTION PAPER

Name of the Examination: WINTER 2022-2023–FAT

Course Code: CSE2007   **Course Title:** Database Management Systems

Set number: 7   **Date of Exam:** 13/05/2023 (FN) (G1)

Duration: 120 Minutes   **Total Marks: 60**

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated

Q1. a. Draw a state diagram and discuss the typical states that a transaction goes through during execution. **(5M)**

     b. Consider the schedule given below

$$S : R_1(X) , R_2(X) , R_3(X), R_1(Y) , R_2(Y) , R_3(Y) , W_1(X) , W_2(Y) W_3(Y)$$

     Check whether the given schedule is conflict-serializable or not. **(5M)**

Q2. Explain various problems and conflicts that arise due to Concurrent Execution with examples. **(10M)**

Q3. a. Describe briefly when a deadlock occurs in Two Phase Locking Protocol with example. **(5M)**

     b. How is Immediate Update better than Deferred Update during Database Recovery. **(5M)**

Q4. Consider the following 2Phase locking protocol. Suppose a transaction T accesses (R&W), a certain set of objects (O1, ..., Ok). The locking happens in the following manner.

1) T acquires exclusive locks to O1, ..., Ok, in increasing order of their addresses.
2) The required read and write operations are performed.
3) All acquired locks are released.

     Justify how does the above scenario guarantee serializability and deadlock freedom. **(10M)**

Q5. Suppose a block holds either 3 data records or 10 search-key/pointer pairs; duplicate search-keys are not allowed. As a function of n (the number of records), how many blocks are needed to hold a data file and dense index, Justify. **(10M)**

Q6. Optimize the following query using heuristic query optimization. **(10M)**
SELECT P.Number, P.Dnum, E.Lname, E.Address, E.Bdate FROM PROJECT AS P, DEPARTMENT AS D, EMPLOYEE AS E WHERE P.Dnum=D.Dnumber AND D.Mgrssn=E.Ssn AND P.Plocation='stafford';

QP MAPPING

Q. No.	Module Number	COMapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	5	5	1,5	-	-	10
Q2	5	5	1,5	-	-	10
Q3	5	5	1,5	-	-	10
Q4	6	6	1,3,5	-	-	10
Q5	6	6	1,3,5	-	-	10
Q6	4	4	1,2,3,5	-	-	10

QUESTION PAPER
Name of the Examination: WINTER 2022-2023 FAT

Course Code: CSE2007
Set number: 8
Duration: 120 Minutes

Course Title: Database Management Systems
Date of Exam: 17/05/2023 (AN) (C2)
Total Marks: 60 Marks

Instructions:
1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. Let us take an example to show the relationship between two FD sets. [10 M]
A relation R (A, B, C, D) having two FD sets
FD1 = {A->B, B->C, A->C} and FD2 = {A->B, B->C, A->D}.

Check whether FD1 and FD2 are equivalent or not? Give the explanation.

Q2. The following transactions show the Two-Phase Locking Techniques for Concurrency Control

T_1	T_2
read_lock(Y);	read_lock(X);
read_item(Y);	read_item(X);
unlock(Y);	unlock(X);
write_lock(X);	write_lock(Y);
read_item(X);	read_item(Y);
$X := X + Y$;	$Y := X + Y$;
write_item(X);	write_item(Y);
unlock(X);	unlock(Y);

Initial values are given as $X=1020$, $Y=1030$
   i) If Serial schedule **T1** followed by **T2** then find the updated **X** and **Y** values? 5 M
   ii) If Serial schedule **T2** followed by **T1** then find the updated **X** and **Y** values? 5 M

Q3. Airline reservation systems (ARS) are systems that allow an airline to sell their inventory (seats). It contains information on schedules and fares and contains a database of reservations (or passenger name records) and of tickets issued (if applicable). ARSs are part of passenger service systems (PSS), which are applications supporting the direct contact with the passenger.

ARS eventually evolved into the computer reservations system (CRS). A computer reservation system is used for the reservations of a particular airline and interfaces with a global distribution system (GDS) which supports travel agencies and other distribution channels in making reservations for most major airlines in a single system.

If you consider booking of a flight ticket as a transaction, How ACID transaction properties will be implemented? Explain. 10 M

Q4. Which recovery concepts can be used for the given scenario in the below diagram using transaction logs? Explain all approaches clearly.

10 M

Checkpoint   **Failure**

t_1 ⊢————————┤

t_2 ⊢———————————┤

t_3 ⊢———————————┤

t_4 ⊢————————┤

Time
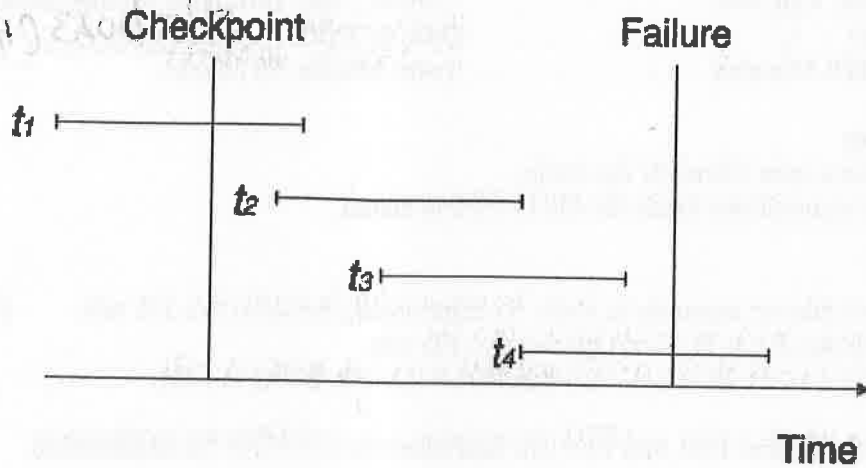
Q5. Why primary index will have less block access than secondary indexing? Justify your answer with a proper example.   10 M

Q6. Suppose that we have an ordered file with r = 150,000 records stored on a disk with block size B = 2,048 bytes. File records are of fixed size and are unspanned, with record length R = 50 bytes.

Find the blocking factor for the file?   3 M
Find the number of blocks needed for the file.   4 M
Calculate the block accesses for a binary search on the data file.   3 M

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	4	1	1	-	10 M
Q2	5	5	1, 5	1, 5	-	10 M
Q3	5	5	1, 5	1, 5	-	10 M
Q4	5	5	1, 5	1, 5	-	10 M
Q5	6	6	1, 3, 5	1, 3, 5	-	10 M
Q6	6	6	1, 3, 5	1, 3, 5	-	10 M

Name of the Examination: WINTER 2022-2023 – FAT

Course Code: CSE2007
Set number: 10

Course Title: Database Management Systems
Date of Exam: 15/05/2023 (AN) (A2)

Duration: 120 Mins

Total Marks: 60

Instructions:

1. Assume data wherever necessary.

2. Any assumptions made should be clearly stated.

Q1. A. Check the given transaction schedule and identify it is satisfies conflict serializability or not. If not, can it be satisfying view serializability? Justify **(6 Marks)**

Time	T1	T2	T3	T4
t1	begin			
t2	read(x)			begin
t3	write(x)			
t4				read(x)
t5	read(z)		begin	
t6			read(y)	
t7			write(y)	
t8		begin		read(y)
t9		read(z)		Commit
t10		write(z)		
t11		commit	read(z)	
t12	write(z)		Commit	
t13	commit			

B. Define serial schedule, non-serial schedule and serializable schedule with example.

(4 marks)

Q2. A. How 2-phase locking protocol overcome the issues related to general locking protocol in database transaction. Add lock and unlock instructions to transactions T_{31} and T_{32}, so that they observe the two-phase locking protocol. **(5 Marks)**

T_{31}: read(A);
   read(B);
   if $A = 0$ then $B := B + 1$;
   write(B).

T_{32}: read(B);
   read(A);
   if $B = 0$ then $A := A + 1$;
   write(A).

B. Check the given transaction and identify which transaction should be Rollback using time stamp ordering protocol. **(5 Marks)**

T1(100)	T2(200)	T3(300)
R(A)		
	R(B)	
W(C)		
		R(B)
R(C)		
	W(B)	
		W(A)

Q3. How does a DBMS create and maintain files of records? How are records arranged on pages, and how are pages organized within a file? Describe each and individual components required for proper transfer of files in a database system. Also describe how does a DBMS keep track of space on disks?

(10 Marks)

Q4. Consider a Hard disk in which Block size = 1000bytes, each record of size 200 bytes. If the total number of records are 10000 and the data entered in Hard disk with and with out any order, what will be the average time complexity to search a record from Hard disk. Also, identify average time complexity to search a record from Index table if Index entry = 20 bytes by considering both sparse and dense indexing.

(10 Marks)

Q5. How the immediate database recovery overcome the limitations of deferred database recovery explain with an example. Write the database log for the following transaction and explain how immediate transaction recovery can apply to that transaction for both undo and redo operation. Consider the initial value of A, B& C= 1000,2000 AND 3000. (10 Marks)

T1	T2
R(A)	
A=A+1000	
W(A)	
R(B)	
B=B+2000	
W(B)	
COMMIT	
	R(C)
	C=C+3000
	W(C)

Q6. Consider the relation schema R(A,B,C,D,E,F) and the functional dependencies A->B,C->DF,AC->E, D->F. What is the primary key of this relation R? What is its highest normal form? Preserving the dependency, decompose R into third normal form. (10 Marks)

QP Mapping

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	5	4,5,6				10
Q2	5	4,5,6				10
Q3	6	5,6				10
Q4	6	5,6				10
Q5	5	4,5,6				10
Q6	3	1,2,3				10

QUESTION PAPER
Name of the Examination: WINTER 2022-2023 – FAT

Course Code: CSE2007
Set number: 12
Duration: 120Mins

Course Title: Database Management Systems
Date of Exam: 17/05/2023 (AN) (CI)
Total Marks: 60

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. A) Given a relation R (A,B,C,D,E,F,G) with the following five functional dependencies **(5M)**

F:
$A \rightarrow BC$
$E \rightarrow CF$
$B \rightarrow E$
$CD \rightarrow EF$
$A \rightarrow G$

Find the closure of each of the attributes and one possible candidate key.

B) Given a relational schema R(M, N, O , P, Q) set of functional dependencies A and B such that:
A = { M $\rightarrow$ N, MN $\rightarrow$ O, P $\rightarrow$ MO, P $\rightarrow$ Q} and B = { M $\rightarrow$ NO, P $\rightarrow$ MQ } using FD sets A and B. **(5M)**

Which of the following option is correct?
i) A is subset of B   ii) B is subset of A   iii) A = B   iv) A ≠ B

Q2. Discuss ACID properties and transaction states of a database transaction with an illustration. **(10M)**

Q3.   **A)** Discuss how serializability is used to enforce concurrency control in a database system. Why is serializability sometimes considered too restrictive as a measure of correctness for schedules? **(5M)**

  B) Consider the below precedence graph. Is the corresponding schedule conflict serializable? Explain your answer. **(5 M)**
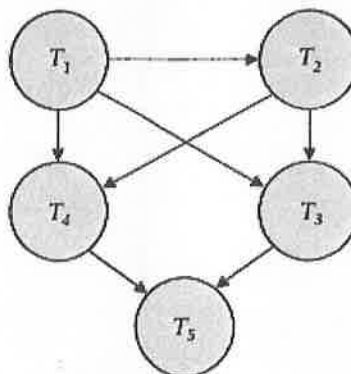
Q4. What are the different types of locks in transaction management? Explain the two-phase locking for concurrency control in detail with suitable examples. **(10M)**

Q5. A) List and explain the physical storage media available on the computers you use routinely. Illustrate the hierarchy of the storage medium based on speed of access. **(5M)**

B) Differenciate fixed-length records and variable-length records. **(5M)**

Q6. Explain why the allocation of records to blocks affects database system performance significantly. Justify your answer. **(10M)**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	3	1,2,3,4	1,4	1	10
Q2	5	5	1,5	1,4	1	10
Q3	5	5	1,5	1,4	1	10
Q4	5	5	1,5	1,4	1	10
Q5	6	6	1,3,5	1,4	1	10
Q6	6	6	1,3,5	1,4	1	10

VIT-AP
UNIVERSITY
Apply Knowledge. Improve Life!®

QUESTION PAPER

Name of the Examination: WINTER 2022-2023 – FAT

Course Code: CSE2007

Course Title: Database Management System

Set number: 013

Date of Exam: 19/05/2023 (ANJ)(E2)

Duration: 120 mins

Total Marks: 60

Instructions:

1. Assume data wherever necessary.

2. Any assumptions made should be clearly stated.

Q1. UPS prides itself on having up-to-date information on the processing and current location of each shipped item. To do this, UPS relies on a company-wide information system. Shipped items are the heart of the UPS product tracking information system. Shipped items can be characterized by item number (unique), weight, dimensions, insurance amount, destination, and final delivery date. Shipped items are received into the UPS system at a single retail centre. Retail centres are characterized by their type, uniqueID, and address. Shipped items make their way to their destination via one or more standard UPS transportation events (i.e., flights, truck deliveries). These transportation events are characterized by a unique scheduleNumber, a type (e.g, flight, truck), and a deliveryRoute.

Create an Entity Relationship diagram that captures this information about the UPS system. Be certain to indicate identifiers and cardinality constraints. **(10M)**

Q2. Four transactions T1, T2, T3, T4 are as follows:

T1={Read(a), Write(a), Read(c), Write(c)};

T2={Read(d), Read(c), Write(d), Read(a) Write(a)};

T3={Read(c), Read(d), Write(d), Read(b), Write(c), Write(b)};

T4={Read(b), Write(b), Read(d), Write(d)}

are submitted at same time and schedule S is made as S=(R1(a), W1(a), R3(c), R3(d), W3(d), R2(d), R2(c), R4(b), W4(b), R1(c), W1(c), R4(d), W4(d),R3(b), W3(c), W3(b), W2(d), R2(a), W2(a)). Find the schedule S is serializable or not through precedence graph and justify your answer. If it is serializable then find the corresponding serial schedule(s). **(10M)**

Q3. Consider the following transactions with data items P and Q initialized to zero:

T1:	T2:
Read(P);	Read(Q);
Read(Q);	Read(P);
If P=0 then	If Q=0 then
Q: =Q+1;	P: =P+1;
Write(Q);	Write(P);

Solve and find any non-serial interleaving of T1 and T2 for concurrent execution leads to a serializable schedule or non-serializable schedule. Explain? **(10M)**

Q4. a) Explain the Check point log-based recovery scheme for recovering the database. **(5M)**

b) Explain the Time Stamp - Based Concurrency Control protocol. How is it used to ensure serializability? **(5M)**

Q5. Consider a file of 8192 records. Each record is 16 bytes long and its key field is of size 6 bytes. The file is ordered on a key field, and the file organization is unspanned (i.e, records are not allowed to cross block boundaries). The file is stored in a file system with block size 512 bytes, and the size of a block pointer is 10 bytes. If the primary index is built on the key field of the file, and a multileveled index scheme is used to store the primary index. Find the number of first level and second level blocks. **(10M)**

Q6. a) Write the working of simple cluster indexing strategy. **(5M)**

b) Table T has 80000 records with record length 50 bytes, block size is 2048 bytes. There are 3000 distinct values for indexing field. How many blocks are required to store the table using simple cluster indexing? **(5M)**

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	2	2	1,2,3,4	-	-	10
Q2	5	5	1,5	-	-	10
Q3	5	5	1,5	-	-	10
Q4	5	5	1,5	-	-	10
Q5	6	6	1,3,5	-	-	10
Q6	6	6	1,3,5	-	-	10

QUESTION PAPER

Name of the Examination: WIN 2022-2023 – FAT

Course Code: CSE2007
Set number: 14
Duration: 120 min

Course Title: Database Management Systems
Date of Exam: 16/05/2023 (AN) (B2)
Total Marks: 60

Answers of all parts of a question should be written together

Answer all six questions.

1. Consider a relation scheme R = (A, B, C, D, E, F) on which the following functional dependencies hold: {A→B, B→A, AC→ E, DE→F, EF→ A}. Check whether this relation is in 3NF or not. If not then normalize to 3NF. **10M**

2. (a) What is a serializable schedule in DBMS? How is it different from a non-serializable schedule? Give an example of a non-serializable schedule.

   (b) Consider the following schedule of transactions:

   T3: R(A)
   T2: R(A)
   T3: W(A)
   T1: R(A) W(A)

   Is the above schedule conflict serializable? If so, provide the equivalent serial schedule. **3+7=10M**

3. Describe dirty read? Consider schedules S1, S2, and S3 below. Determine whether each schedule is strict, cascadeless, recoverable, or nonrecoverable. (ri → read by transaction Ti, wi → write by transaction Ti)

   S1: r1 (X); r2 (Z); r1 (Z); r3 (X); r3 (Y); w1 (X); c1; w3 (Y); c3; r2 (Y); w2 (Z); w2 (Y); c2;
   S2: r1 (X); r2 (Z); r1 (Z); r3 (X); r3 (Y); w1 (X); w3 (Y); r2 (Y); w2 (Z); w2 (Y); c1; c2; c3;
   S3: r1 (X); r2 (Z); r3 (X); r1 (Z); r2 (Y); r3 (Y); w1 (X); c1; w2 (Z); w3 (Y); w2 (Y); c3; c2;

   10M

4. (a) How does the two-phase locking protocol ensure serializability?

   (b) Suppose a schedule S1 is executing some operations on a database using two transactions T0 and T1. Now, describe all the recovery actions will be taken place for the following two logs using immediate database modification recovery.

I. <T0, Start>	II. <T0, Start>
<T0, X, 100, 200>	<T0, X, 100, 200>
<T0, Y, 300, 250>	<T0, Y, 300, 250>
<T0, commit>	<T0, commit>
<T1, Start>	<T1, Start>
<T1, Z, 700. 600>	<T1, Z, 700. 600>
	<T1, commit>

   5+5=10M

5. (a) What are the advantages of indexing in DBMS?

   (b) Differentiate between prime indexing and secondary indexing. **5+5=10M**

6. (a) Consider a table with 1,000,000 records and a secondary index on a key (or non-key) field with an average length of 30 characters. Each index entry requires 50 bytes of space. If the block size is 8 KB, how many blocks are required to store the index file? Calculate number of block access to read a value from database table.

(b) Consider the example in question 4.a is following primary index with search key of 3-byte string and block pointer size 4 bytes. Calculate the number of block access to search a value.

5+5=10M

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	4	4	PO1, PO2, PO3, PO5			10
Q2	5	5	PO1, PO5			10
Q3	5	5	PO1, PO5			10
Q4	5	5	PO1, PO5			10
Q5	6	6	PO1, PO3, PO5			10
Q6	6	6	PO1, PO3, PO5			10

VIT-AP
UNIVERSITY
Apply Knowledge. Improve Life!

QUESTION PAPER

Name of the Examination: FAT(SHORT SUMMER_II)

Course Code: CSE2007

Set number: 1

Duration:120 Min.

Course Title: Database Management Systems

Date of Exam: 21/07/2023 (FN) (C1)

Total Marks: 60

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1(a). How can we reconcile the need for data independence with the increasing inter connectivity and interdependence of various data sources and systems in today's complex technological landscape? **6M**

Q1(b) consider the diagram below.explain in detail about strong entity type and weak entity type and how a relationship is performed between these two. **6M**
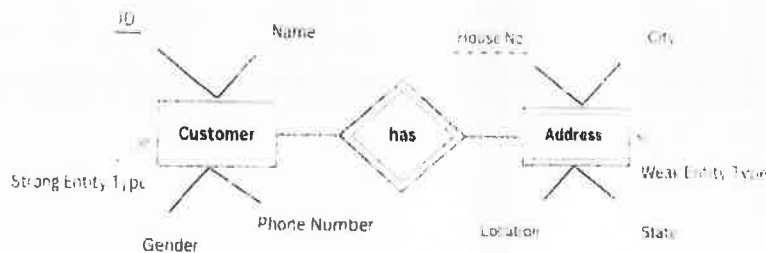
Q2. What are the different types of constraints in a relational database system, and how do they ensure data integrity and enforce business rules during data manipulation operations? Explain each with example by taking Employee Relation. **12M**

Q3. what is functional dependency? Given a relation R(A, B, C, D) and Functional Dependency set FD = { AB → CD, B → C }, determine whether the given R is in 2NF? If not convert it into 2 NF.

12M

Q4. create an employee table with the attributes of EmpID,Emp_Name,Salary.what are the general steps involved in query processing explain each with the Example query: SELECT * FROM Employees WHERE Salary > 50000 **12M**

Q5. What are the commonly used indexing Techniques in DBMS to improve query performance by allowing efficient data retrieval based on specific attributes.explain each technique by taking example.

12M

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	1	1	1	2	-	12
Q2	2	1	1	2	-	12
Q3	3	1	1	2	-	12
Q4	4	2	4	1	-	12
Q5	6	2	4	1	-	12

VIT-AP
UNIVERSITY
Apply Knowledge. Improve Life!®

QUESTION PAPER

Name of the Examination: Fast Track Fall 2023-24 Semester – FAT

Course Code: CSE2007   **Course Title:** Database Management Systems

Set number: 1   **Date of Exam:** 22/8/23 (E2)(AN)

Duration: 120Min   **Total Marks:** 60

ClassID-AP2023241000466

Instructions:

1. Assume data wherever necessary.
2. Any assumptions made should be clearly stated.

Q1. Consider a relation R (A, B, C, D, E, F) and F={A->BC, F->A, C->A, D->E, E->DA} and the relation is decomposed into three new relations R1(ACD), R2(BCD), R3(EFD) verify the following

   a) Whether it is lossy or lossless join b) Dependency preservation or not. **(12M)**

Q2. Consider the following SQL query on the BOOK database and transform these SQL queries into relational algebra expressions.

Draw the initial query tree for the expression and then drive their optimized query tree after applying heuristic rules on them.

QUERY: select ISBN, Book_title, year, page_count from AUTHOT A, Book B, Author_Book AB where B.ISBN=AB.ISBN and AB.AID=A.AID and Aname='Neha';   **(12M)**

Q3. Check the given Schedule is view serializable or not, if yes then give the serial schedule.

T1	T2	T3
R (A)		
	R (A)	
		W (A)
W (A)		

(12M)

Q4. a) Apply Time stamp ordering protocol and identify which transaction is roll back in a given schedule.   (6M)

T1	T2	T3
R(A)		
	R(B)	
W (C)		
		R(B)
R (C)		
	W(B)	
		W(A)

b) What is the need of concurrency control? What are the problems exists in concurrent access. **(6M)**

Q5.   a) A student file with Rollno as the key field includes records with the following rollno values. 7, 10, 1, 23, 5, 15, 17, 19, 11, 39, 35, 8, 40, 25 Suppose that the search field values are inserted in the given order in a B+ tree of order p=5, show the tree after inserting these values.   **(6M)**

b)   A PARTS file with part# as the hash key includes records with the following Part# values: 2369, 3760, 4692, 4871, 5659, 1821, 1074, 7115, 1620, 2428, 3943, 4750, 6975, 4981, and 9208. The file uses 8 buckets numbered from 0 to 7. Each bucket is one disk block and holds two records. Load these records using the hash function H(K)= K mod 8. Calculate the average number of block accesses for a random retrieval on Part#.

(6M)

QP MAPPING

Q. No.	Module Number	CO Mapped	PO Mapped	PEO Mapped	PSO Mapped	Marks
Q1	3	3	1,2,3,4	2	1	12
Q2	4	4	1,2,3,5	2	1	12
Q3	5	5	1,5	2	1	12
Q4	5	5	1,5	2	1	12
Q5	6	6	1,3,5	2	1	12