Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Fourth Semester B.Tech Degree Examination July 2021 (2019 Scheme)

Course Code: CST204
Course Name: DATABASE MANAGEMENT SYSTEMS

Max. Marks: 100 Duration: 3 Hours

PART A

(Answer all questions; each question carries 3 marks) Marks

1	List any three categories of database users, highlighting any one important characteristic of each category.	3
2	What are the major difference between structured, unstructured and semi structured data.	3
3	What is entity integrity? Why is it important?	3
4	Distinguish between Super key, Candidate key, and Primary key using a real convincing example.	3
5	Illustrate the concept of trigger in SQL with an example	3
6	Compare DDL and DML with the help of an example	3
7	Illustrate different anomalies in designing a database	3
8	How can we conclude two FDs are equivalent?	3
9	Illustrate two phase locking	3
10	How conversions of locks are achieved in concurrency control?	3

PART B

(Answer one full question from each module, each question carries 14 marks)

Module -1

11	a)	A company has the following scenario: There are a set of salespersons. Some of them manage other salespersons. However, a salesperson cannot have more than one manager. A salesperson can be an agent for many customers. A customer is managed by exactly one salesperson. A customer can place any number of orders. An order can be placed by exactly one customer. Each order lists one or more items. An item may be listed in many orders. An item is assembled from different parts and parts can be common for many items. One or more employees assemble an item from parts. A supplier can supply	7

different parts in certain quantities. A part may be supplied by different suppliers.

(i) Identify and list entities, suitable attributes, primary keys, and relationships to represent the scenario.

(ii) Draw an ER diagram to model the scenario using min-max notation.

b) Explain three schema architecture with figure 7

12 a) Illustrate Database architecture with a neat diagram 7

b) Explain the characteristics of Database system 7

Module -2

13 a) Study the tables given below and write relational algebra expressions for the queries that follow. 8

STUDENT(ROLLNO, NAME, AGE, GENDER, ADDRESS, ADVISOR)

COURSE(COURSEID, CNAME, CREDITS)

PROFESSOR(PROFID,PNAME, PHONE)

ENROLLMENT(ROLLNO, COURSEID, GRADE)

Primary keys are underlined. ADVISOR is a foreign key referring to PROFESSOR table. ROLLNO and COURSEID in ENROLLMENT are also foreign keys referring to the primary keys with the same name.

(i) Names of female students

(ii) Names of male students along with adviser name

(iii) Roll Number and name of students who have not enrolled for any course.

b) Explain the left outer join, right outer join, full outer join operations with examples 6

14 a) Consider the following relations for a database that keeps track of business trips of salespersons in a sales office: 9

SALESPERSON(Ssn, Name, StartYear, DeptNo)

TRIP(Ssn, FromCity, ToCity, DepartureDate, ReturnDate, TripId)

EXPENSE(TripId, AccountNo, Amount)

i) A trip can be charged to one or more accounts. Specify the foreign keys for this schema, stating any assumptions you make.

ii) Write relation algebra expression to get the details of salespersons who have travelled between Mumbai and Delhi and the travel expense is greater that Rs. 50000.

iii) Write relation algebra expression to get the details of salesperson who had incurred the greatest travel expenses among all travels made.

b) List the basic data types available for defining attributes in SQL? 5

Module -3

15 a) Illustrate structure of B-Tree and B+ Tree and explain how they are different? 5

b) What are the different types of single-level ordered indices? Explain. 9

16 a) Differentiate between static hashing and dynamic hashing. 9

b) Write short notes on Nested queries 5

Module -4

17 a) i) What are Armstrong's axioms? 10

ii) Write an algorithm to compute the attribute closure of a set of attributes (X) under a set of functional dependencies (F).

iii) Explain three uses of attribute closure algorithm.

b) Explain the difference between BCNF and 3NF with an example 4

18 a) Consider the relation R = {A, B, C, D, E, F, G, H} and the set of functional dependencies F = {A→DE, B→F, AB→C, C→GH, G→H}. What is the key for R? Decompose R into 2NF and then 3NF relations. 9

b) What is the lossless join property of decomposition? Why is it important? 5

Module -5

19 a) Explain the concepts behind the following: - 10

i) Log-Based Recovery

ii) Deferred Database Modification.

b) Why recovery is needed in transaction processing? 4

20 a) Differentiate serial and concurrent schedules. Elaborate conflict serializability with suitable example. 10

b) What are the desirable properties of transactions? Explain 4

*****

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Fourth Semester B.Tech Degree Examination June 2022 (2019 scheme)

Course Code: CST204

Course Name: DATABASE MANAGEMENT SYSTEMS

Max. Marks: 100 Duration: 3 Hours

PART A

(Answer all questions; each question carries 3 marks) Marks

1		List any three characteristics of database system	3
2		Draw neat labelled diagram of three schema architecture and briefly describe each level	3
3		Write briefly about any three relational database integrity constraints.	3
4		Differentiate between theta join and natural join operations.	3
5		Give any three uses of a trigger	3
6		A file has r =20000 STUDENT records of fixed length. Each record has the following fields: NAME (30 bytes), SSN (9 bytes), ADDRESS (40 bytes), PHONE(9 bytes), BIRTHDATE (8 bytes), GENDER (1 byte), DEPTID (4 bytes), CLASSCODE (4 bytes), and PROGID (3 bytes). An additional byte is used as a deletion marker. The file is stored on the disk with block size B=512 bytes,	3
	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 average time it takes to find a record by doing a linear search	
7		Define Boyce-Codd normal form. How does it differ from 3NF?	3
8		Suppose, a relational schema R (P,Q, R, S) and set of functional dependencies F and G are as follow: F : { P → Q, Q → R, R → S } G : { P → QR, R → S }. Check the equivalency of functional dependencies F and G.	3
9		Write briefly on log based recovery	3
10		Explain briefly the characteristics of Column family database.	3

PART B

(Answer one full question from each module, each question carries 14 marks)

Module -1

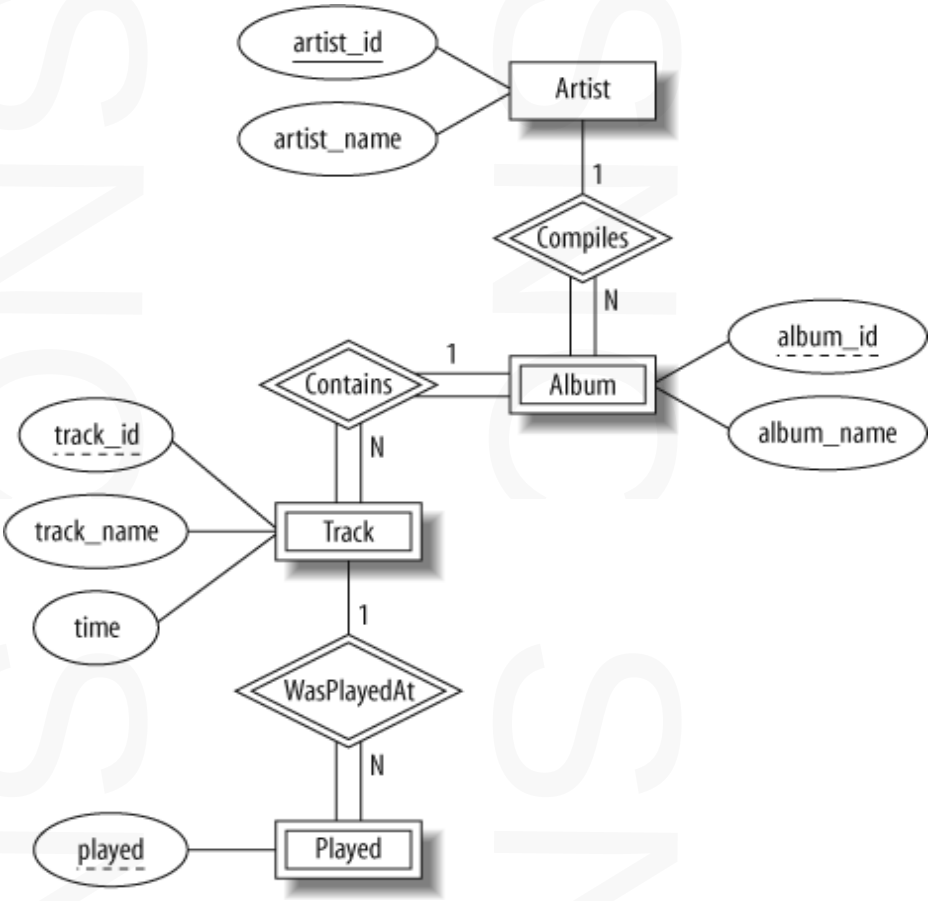
11	a)	Differentiate between two-tier and three-tier client-server database architecture with the help of neat labelled diagrams.	7

b) Draw an ER diagram based on the following information, 7

- Manufacturers have a name, which we may assume is unique, an address, and a phone number
- Products have a model number and a type. Each product is made by one manufacturer, and different manufacturers may have different products with the same model number. However, you may assume that no manufacturer would have two products with the same model number
- Customers are identified by their unique social security number. They have email addresses, and physical addresses. Several customers may live at the same (physical) address, but we assume that no two customers have the same email address
- An order has a unique order number, and a date. An order is placed by one customer. For each order, there are one or more products ordered, and there is a quantity for each product on the order.

12 a) Write briefly about any three types of database end users 6

b) Interpret the following ER diagram 8



Module -2

13 a) Consider the following schema, 8

Suppliers (sid , sname, address)

Parts (pid, pname, color)

Catalog (sid, pid, cost)

The primary key fields are underlined.

Write relational algebra expressions for the following queries:

b) Find the name of parts supplied by supplier with sid=105

ii) Find the names of suppliers supplying some green part for less than Rs 1000

iii) Find the IDs of suppliers who supply some red or green part

iv) Find the names of suppliers who supply some red part

b) Differentiate between the following SQL statements 6

b) DROP and DELETE

ii) ALTER and UPDATE

14 a) Write SQL DDL statements based on the following database schema (Assume suitable domain types): 8

Employee (eid, name, designation, salary, comp_id)

Company (comp_id, cname, address, turnover)

b) Create the above mentioned tables assuming each company has many employees. Mention the primary key, foreign key and not null constraints.

ii) Insert values into both the tables. Mention in which order insertions will be carried out.

ii) Modify the table Employee to include a new column "years_of_exp"

iv) Increment the salary of employees whose salary is less than Rs25000 by 5%

b) Illustrate any three ways of using INSERT statement in SQL. 6

Module -3

15 a) For the relation schema below, give an expression in SQL for each of the queries that follows: 8

employee (ID, person_name, street, city)

works (ID, company_name, salary)

company (company_name, city)

manages (ID, manager_id)

b) Find the employees whose name starts with 'C'

ii) Find the name of managers of each company

iii) Find the ID, name, and city of residence of employees who works for "First Bank Corporation" and earns more than Rs50000

iv) Find the name of companies whose employees earn a higher salary, on average, than the average salary at "First Bank Corporation"

b) Differentiate correlated and non-correlated nested queries with suitable examples 6

16 a) What is multi-level indexing? How does it improve the efficiency of searching an index file? 8

b) Insert the following keys, in the order given, into a B -tree of order 3: {10, 50, 20, 5, 22, 25} 6

Module -4

17 a) Consider a relation R(A, B, C, D, E) with FDs 8

AB → C, AC → B, BC → A, D → E.

Determine all the keys of relation R. Also decompose the relation into collections of relations that are in BCNF.

b) Write briefly on the different types of anomalies in designing a database. 6

18 a) Consider a relation schema R (A,B,C,D) with the following functional dependencies A → B, B → C, C → D, D → B. Determine whether the decomposition of R into R1 (A , B) , R2 (B , C) and R3 (B , D) is lossless or lossy. Write the complete steps. 6

b) What is dependency preservation property for decomposition? Why is it important? 8

Module -5

19 a) Explain briefly the ACID properties of a transaction. 8

b) Check whether the given schedules are conflict serializable or not 6

i) S1 : $R_1(X)$, $R_2(X)$, $R_1(Y)$, $R_2(Y)$, $R_3(Y)$, $W_1(X)$, $W_2(Y)$

ii) S2 : $R_1(X)$, $R_2(X)$, $R_2(Y)$, $W_2(Y)$, $R_1(Y)$, $W_1(X)$

20 a) What is two phase locking protocol? How does it guarantee serializability? 8

b) What are the main characteristics of NOSQL systems in the areas related to data models and query languages? 6

******

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Fourth Semester B.Tech Degree Supplementary Examination June 2023 (2019 scheme)

Course Code: CST 204

Course Name: DATABASE MANAGEMENT SYSTEMS

Max. Marks: 100 Duration: 3 Hours

PART A

(Answer all questions; each question carries 3 marks) Marks

1	Classify the following cases into logical data independence and physical data independence. (a) Creating an index for a data file (b) Changing the integrity constraint (c) Reorganizing the file	3
2	What is meant by a recursive relationship type? Give an example of recursive relationship type.	3
3	Consider the relational model constraints: domain constraint, key constraint, entity integrity and referential integrity. Specify which of these constraints may be violated during the following modification operations: insert, update and delete.	3
4	What is meant by complete set of relational algebra? Show how join operation in relational algebra can be represented using this set.	3
5	What is meant by a correlated nested query? Give a suitable example.	3
6	Explain the advantage of a multilevel index.	3
7	Why Armstrong's axioms are said to sound and complete?	3
8	What is meant by lossless join property?	3
9	List and explain the desirable properties of a transaction.	3
10	Illustrate the states for transaction execution.	3

PART B

(Answer one full question from each module, each question carries 14 marks)

Module -1

11	Consider the following information about a university database: Professors have a ssn, a name, an age, a rank, and a research specialty. Projects have a project number, a sponsor name, a starting date, an ending date, and a budget. Graduate students have ssn, a name, an age, and a degree program (e.g., M.S. or Ph.D.). Each project is managed by one professor (known as the project's principal investigator). Each project is worked on by one or more professors (known as the project's co-investigators). Professors can manage and/or work on multiple projects. Each project is worked on by one or more graduate students (known as the project's research assistants). When graduate students work on a project, a professor must supervise their work on the project. Departments have a department number, a department name, and a location. Departments have a professor (known as the chairman) who manages the department. Professors work in one or more departments, and for each department that they work in, a time	14

percentage is associated with their job. Graduate students have one major department in which they are working on their degree. Each graduate student has another, more senior graduate student (known as a student advisor) who advises him or her on what courses to take. Design and draw an ER diagram that captures the information about the university.

12 a) Explain the difference between database schema and database state with suitable example. Specify the role of schema in a DBMS. 7

b) With a neat diagram explain the Three Schema Architecture of a DBMS 7

Module -2

13 a) Consider the two relations T1 and T2 shown below. Show the results of the following operations. 8

Relation **T1**

P	Q	R
30	Ac	25
35	Bc	28
45	Ac	26

Relation **T2**

A	B	C
30	Bc	26
45	Cc	23
30	Bc	25

i) $\mathbf{T1} \bowtie_{T1.Q=T2.B} \mathbf{T2}$

ii) $\mathbf{T1} \bowtie_{T1.P=T2.A} \mathbf{T2}$

iii) $\mathbf{T1} \cup \mathbf{T2}$

iv) $\mathbf{T1} \bowtie_{(T1.P=T2.A \text{ AND } T1.R=T2.C)} \mathbf{T2}$

b) An Employee relation has attributes: Employee-Id (numeric type), Name (character type), Salary (numeric type) and Dep-No (numeric type). 6

A Department relation has attributes: Department-Number (numeric type), Department-Name(character type), Dep-Manager-Id (numeric type).

Employee-Id is the primary key of Employee relation. Department-Number is the primary key of the Department relation. Dep-No attribute of Employee relation refers to the Department-Number attribute of Department relation and Dep-Manager-Id attribute of Department relation refers to the Employee-Id attribute of Employee relation.

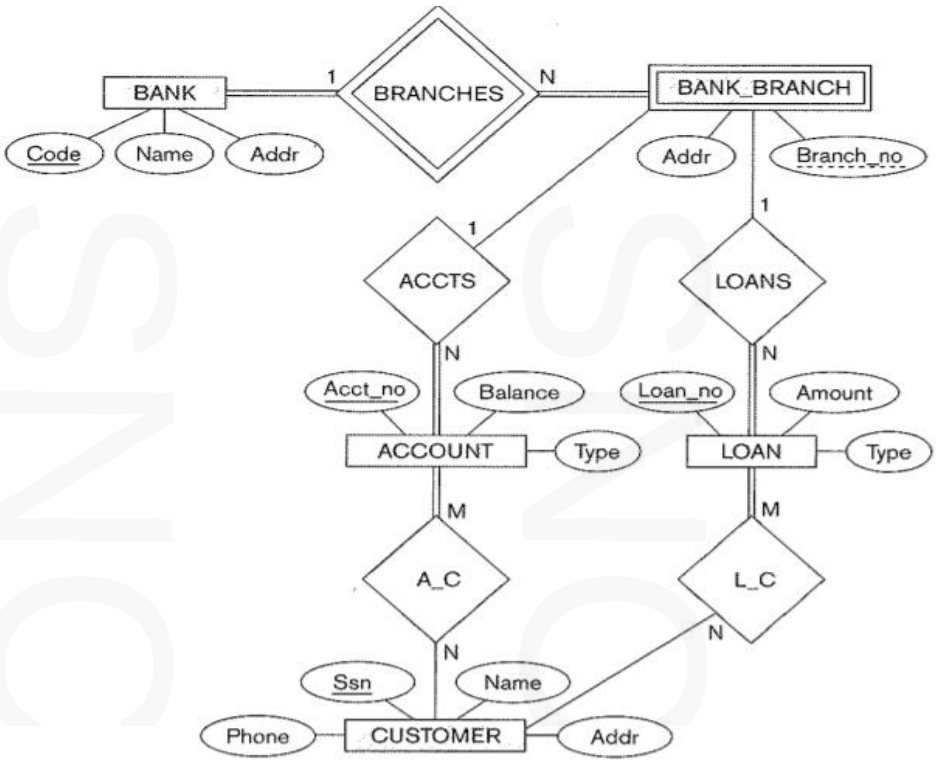
(i) Write create table statements by specifying necessary integrity constraints for creating these two relations in SQL.

(ii) Write SQL statement to insert the details of an employee John with id 101 with salary 5000 and working in department number 5.

(iii) Insert the details of a Research Department with Department Number 1 and it has not been assigned any manager.

(iv) Assume that a department with employees working in it is to be deleted. Specify the two options to manage this scenario.

14 Convert the ER schema for Bank database given below into a relational schema. Specify all primary keys and foreign keys. 14



Module -3

15 a) Consider the following relations: 7

Employee (Employee-Id, Employee-Name, Salary, Department-No)
Department (Department-No, Department-Name)
Write SQL queries for the following:

(i) Retrieve the employee names and their department names
(ii) Retrieve department names and the average salary given by them
(iii) Retrieve the ids of employees getting salary greater than the average salary of their department
(iv) For each department that has more than 4 employees, retrieve the department-No and the number of employees getting salary more than Rs. 50000

b) What is meant by a heap file? Explain how insert, update, delete and search operations can be performed in a heap file. 7

16 a) What are the advantages of Views? Explain two view implementation techniques. 7

b) Consider a disk with block size 512 bytes. A block pointer is 6 bytes long, and a record pointer is 7 bytes long. A file has 30,000 EMPLOYEE records of fixed-length. Each record has the following fields: NAME (30 bytes), SSN (9bytes), DEPARTMENTCODE (9 bytes), ADDRESS (40 bytes), PHONE (9 bytes),BIRTHDATE (8 bytes), SEX (1 byte), JOBCODE (4 bytes), SALARY (4 bytes, real number). An additional byte is used as a deletion marker. Assume that file is not ordered by the key field SSN and we need to create a secondary index on SSN. 7

(i) Find the number of levels needed, if we make it into a multilevel index.
(ii) Find the number of block accesses needed to retrieve a record from this file if we use the multilevel index.

Module -4

17 a) Given relation R(A,B,C,D,E) and functional dependencies F={AB → C, CE → D, A → E}. Determine whether each functional dependency below is in $\mathbf{F}^+$ or not: 6

i) AB → D

ii) A → C

b) Consider the following relation: 8

CAR_SALE(Car#, Date_sold, Salesperson#, Commission%, Discount_amt)

Assume that a car may be sold by multiple salespeople, and hence {Car#,Salesperson#} is the primary key. Additional dependencies are :

Date_sold → Discount_amt and Salesperson# → Commission%

(i) Based on the given primary key and functional dependencies, is this relation in 1NF, 2NF, or 3NF? Why or why not?

(ii) How would you successively normalize it completely?

18 a) Consider the following decompositions for the relation schema R into R1, R2 and R3. Determine whether the decomposition has the lossless join property with respect to the given F. 7

R={P, Q, R, S, T, U}

R1={P, Q}, R2= {R, S, T}, R3={P,R,U}

F = { P → Q, R → {S, T}, {P,R} → U}

b) Explain insert, update and delete anomalies with suitable examples. 7

Module -5

19 a) Consider the schedule S of three transactions T1,T2 and T3 given below. State whether the schedule is serializable or not. 7

S: r3(Y), r3(Z), r1(X),w1(X), w3(Y), w3(Z),r2(Z), r1(Y), w1(Y), r2(Y), w2(Y), r2(X), w2(X)

(Hint: Interpret the notation r3(Y) as the operation read database item Y of transaction T3.)

b) Explain the lost update problem and temporary update problem that occur when concurrent execution is uncontrolled. 7

20 a) Explain conservative and strict two-phase locking techniques. Why strict 2PL is deadlock free? 7

b) Differentiate among recoverable, cascading rollback and strict schedules with suitable examples. 7

***

Reg No.:_____ Name:_____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

B.Tech Degree S4 (R,S) / S4 (PT) (R,S) Examination June 2023 (2019 Scheme)

Course Code: CST 204

Course Name: Database Management Systems

Max. Marks: 100 Duration: 3 Hours

PART A

(Answer all questions; each question carries 3 marks) Marks

1 List any SIX major advantages of using a DBMS 3

2 What is the concept of a weak entity used in data modelling? Define the terms owner entity type, Identifying relationship type. 3

3 Define theta join. 3

Given the two relations R and S:

A	B	C
1	2	3
4	5	6
7	8	9

D	E
3	1
6	2

Find $R\theta_{B<D}S$.

4 Define primary key, candidate key and super key. 3

5 What is the difference between the WHERE and HAVING clause? Illustrate with an example. 3

6 Explain the difference between Hash indexes and B+-tree indexes. 3

7 Define the term functional dependency. Why are some functional dependencies called trivial?

8 List Armstrong Axiom rules 3

9 List the ACID properties of transactions. 3

10 What is a key-value database? List its major properties. 3

PART B

(Answer one full question from each module, each question carries 14 marks)

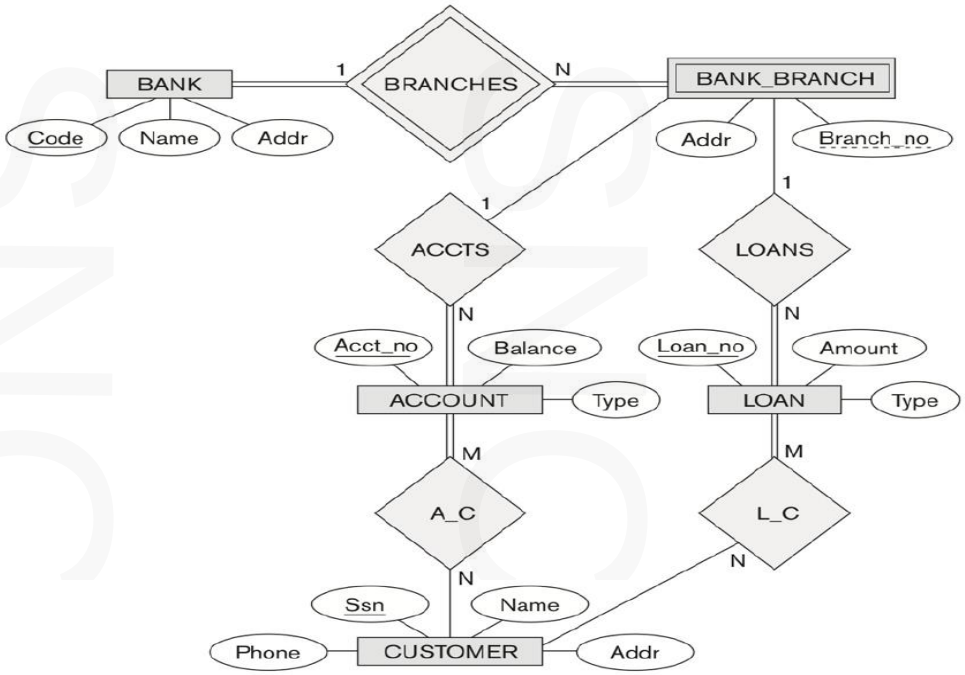
Module -1

11 Draw an ER diagram to model the application with the following assumptions. Specify key attributes of each entity type and (min, max) constraints on each relationship type. 14

- Each home uniquely defined by home identifier, street address, city, state, a number of bedrooms and a number of bathrooms and an associated owner.
- Each owner has a Social Security Number, first name, last name, phone, and profession.
- An owner can spouse one or more homes.
- Agents represent owners in the sale of a home. An agent can list many homes, but only one agent can list a home.
- An agent has a unique agent number, name, phone number and an associated office.
- When an owner agrees to list a home with an agent, a commission and a selling price are determined.
- An office has office identifier, phone number, the manager name, address and an optional agent number.
- Many agents can work at one office.
- A buyer entity type has a Social Security Number, first name, last name, phone, preferences for the number of bedrooms and bathrooms, and a price range.
- An agent can work with many buyers, but a buyer works with only one agent.

12 a) What is the difference between logical data independence and physical data independence? Which one is harder to achieve? Why? 6

b) 8



Consider the bank database given above and answer the following questions

i. List the strong (nonweak) 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 constraints do the partial key and the identifying relationship of the weak entity type specify in this diagram?
iv. List the names of all relationship types, and specify the (min, max) constraint on each participation of an entity type in a relationship type.

v. 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 does this show up on the (min, max) constraints?

Module -2

13 a) Consider the UNIVERSITY database with the following relations: 10

STUDENT (**rollNo**, name, degree, year, sex, deptNo, advisor)

DEPARTMENT (**deptId**, name, hod, phone)

PROFESSOR (**empId**, name, sex, startYear, deptNo, phone)

COURSE (**courseId**, cname, credits, deptNo)

ENROLLMENT (**rollNo, courseId, sem, year**, grade)

TEACHING (**empId, courseId, sem, year**, classRoom)

PREREQUISITE(**preReqCourse, courseID**)

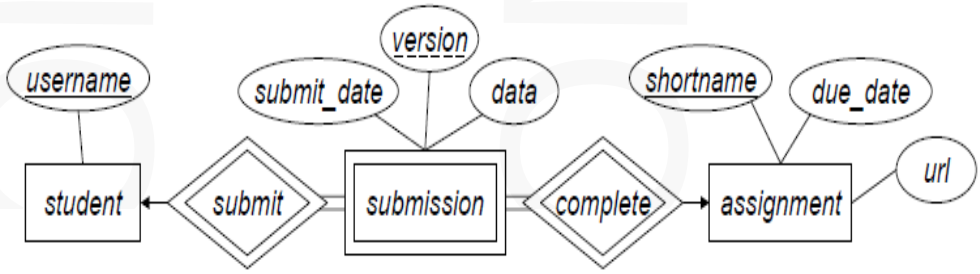
Write relational algebra expressions for the following queries:

i. For each department, find its name and the name, sex and phone number of the head of the department.
ii. Find courses offered by each department.
iii. Find those students who have registered for all courses offered in the department of Computer Science.
iv. Obtain the department Ids for departments with no lady professor.
v. Obtain the rollNo of girl students who have obtained at least one S grade.

b) What is a foreign key constraint? Why are such constraints important? 4

What is referential integrity?

14 a) Convert the following ER diagram into a relational schema 4



b) Consider the following relation schema with referential integrity constraints: 10

STUDENT (**rollNo**, name, degree, year, sex, deptNo, advisor)

DEPARTMENT (**deptId**, name, hod, phone)

PROFESSOR (**empId,** name, sex, startYear, deptNo, phone)

Write SQL DDL statements for the following:

i. Create table STUDENT, DEPARTMENT, PROFESSOR including primary and foreign key integrity constraints.

ii. Add an address attribute in the table STUDENT

iii. Write an SQL statement to delete the "CS" department. Given the referential integrity constraints, explain what happens when this statement is executed.

Module -3

15 a) What is an assertion? How they differ from triggers? 4

b) Consider the following relation schema and write SQL queries to find: 10

EMPLOYEE(Fname, Minit, Lname, **SSN**, Bdate, Address, Sex, Salary SuperSSN, Dno)
DEPARTMENT(Dname, **Dnumber**, MgrSSN, MgrStartDate)
DEPT_LOCATIONS(**Dnumber, Dlocaions**)
PROJECT(Pname, **Pnumber,** Plocation,Dnum)
WORKS_ON(**ESSN, Pno**,Hours)

i. Retrieve the name and address of all employees who work for the 'Research' department.

ii. For each employee, retrieve the employee's name, and the name of his or her immediate supervisor.

iii. Retrieve the name of each employee who works on all the projects controlled by department number 5.

iv. Make a list of all project numbers for projects that involve an employee whose last name is 'Smith' as a worker or as a manager of the department that controls the project.

v. Retrieve the SSN of all employees who work on project number 1, 2, or 3.

16 a) Consider a disk with block size B =512 bytes. A block pointer is P=6 bytes long and a record pointer is P_R =7 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. 9

i. Calculate the record size R in bytes.

ii. Suppose that the file is ordered by the key field Ssn and we want to construct a primary index on Ssn. Calculate The number of first-level index entries and the number of first-level index blocks

iii. Calculate the number of levels needed if we make it into a multilevel index.

b) What is a grid file? What are its advantages and disadvantages? 5

Module -4

17 a) Consider a relation R with five attributes (A,B,C,D,E) . You are given the following dependencies: A → B, BC → E, and ED → A. 6

i. List all keys for R.

ii. Is R in 3NF?

iii. Is R in BCNF?

b) Define minimal cover. Let the given set of functional dependencies be: $E: \{B \rightarrow A, D \rightarrow A, AB \rightarrow D\}$. Find the minimal cover of E 8

18 a) Explain with example 2NF, 3NF and BCNF. 8

b) Consider a relation schema R(X Y Z W P) (above table R) is decomposed into R1(X Y Z) and R2(Z W P). Determine whether the above R1 and R2 are Lossless or Lossy? 6

Module -5

19 a) What is a schedule? Define the concepts of recoverable, cascade less and strict schedules, and compare them in terms of their recoverability. 8

b) Which of the following schedule is conflict serializable? For each serializable schedule determine the equivalent serial schedule. 6

(a) r1 (X); r3 (X); w1(X); r2(X); w3(X)
(b) r1 (X); r3 (X); w3(X); w1(X); r2(X)
(c) r3 (X); r2 (X); w3(X); r1(X); w1(X)

20 a) What is the two-phase locking (2PL) protocol? How does it guarantee serializability? How strict 2PL differs from basic 2PL? 7

b) Explain the need for multimodal database. List the important characteristics of ArangoDB. 7

***