

Reg. No: _____

Name: _____

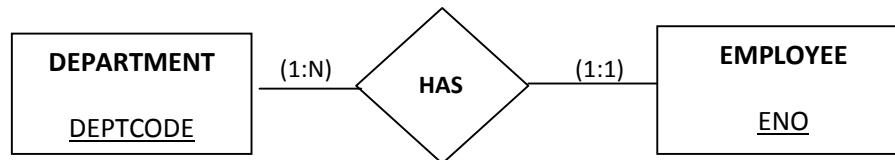
APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
FOURTH SEMESTER B.TECH DEGREE EXAMINATION, JUNE 2017
CS208: PRINCIPLES OF DATABASE DESIGN (CS, IT)

Max. Marks: 100

Duration: 3 hours

Limit answers to the required points.**PART A****Answer all questions**

1. List out any *three* salient features of database systems. (3)
2. How is DML different from DDL? Write a sample statement in DML and one in DDL. (3)
3. Can we represent the situation modelled by the following ER diagram without the relationship 'HAS'? If so, draw the new diagram. If not, give the reasons. (Entities are DEPARTMENT and EMPLOYEE. Attributes names are given under entity names; keys are underlined.) (3)



4. Consider the a relation R(A,B,C,D) where A is a key of R. Write any three relational algebra expressions equivalent to $\prod_{A,B} (\sigma_{A=2 \text{ and } B=3} (R))$ (3)

PART B**Answer any two full questions**

5. Study the tables given below and write relational algebra expressions for the queries that follow. (9)

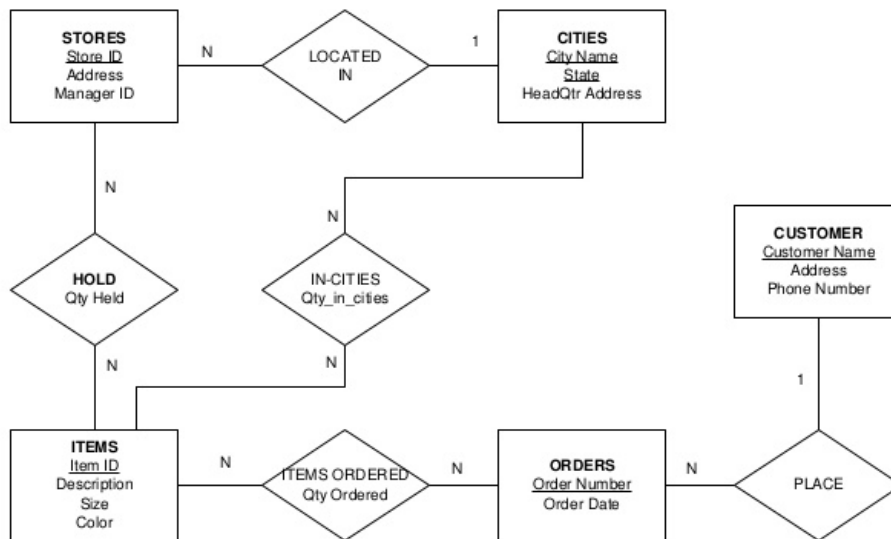
STUDENT(ROLLNO, NAME, AGE, GENDER, ADDRESS, ADVISOR)COURSE(COURSEID, CNAME, CREDITS)PROFESSOR(PROFID, PNAME, PHONE)ENROLLMENT(ROLLNO, COURSEID, GRADE)

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.

6. 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 different parts in certain quantities. A part may be supplied by different suppliers. (9)
- (i) Identify and list entities, suitable attributes, primary keys, foreign keys and relationships to represent the scenario.
 - (ii) Draw an ER diagram to model the scenario using *min-max* notation.
7. a. Justify the importance of weak entity sets with the help of an example. (3)
- b. In the ER diagram below, names of entity sets and relationships are shown in capital and corresponding attributes are listed under each such name. Key attributes are underlined. All the *participations* are *total*. Use the standard synthesis procedure to convert the ER diagram into the corresponding relational schema. Clearly show primary and foreign keys. (6)



PART C

Answer all questions

8. Illustrate the GROUP BY clause with the help of a real example. (3)
9. Consider the query SELECT NAME, AGE FROM STUDENT WHERE GENDER = 'Male' on the table STUDENT(ROLLNO, NAME, AGE, GENDER, ADDRESS). Give a relational algebra expression corresponding to the query. Is result produced by the query and your expression always the same? Why? (3)
10. Determine any two candidate keys of the relation R(A,B,C,D,E,F) with FDs AB→C, C→AD, D→EF, F→B. (3)
11. Give an example for a relation that has insertion, deletion and update anomalies. Which type(s) of functional dependency can formally model these anomalies? Quote one such (3)

dependency from your example

PART D

Answer any two full questions

12. a. Illustrate the use of assertions with a typical example. (3)
- b. Consider a relation (A,B,C,D,E,F) with A as the only key. Assume that the dependencies $E \rightarrow F$ and $C \rightarrow DEH$ hold on R. (6)
- (i) Is R is in 2NF? If not, decompose to 2NF.
- (ii) Is R is in 3NF? If not, decompose to 3NF.
13. In the following tables ADVISOR and TAUGHTBY are foreign keys referring to the table PROFESSOR. ROLLNO and COURSEID in ENROLLMENT refer to tables with primary keys of the same name. (9)
- STUDENT(ROLLNO, NAME, AGE, GENDER, ADDRESS, ADVISOR)
- COURSE(COURSEID, CNAME, TAUGHTBY, CREDITS)
- PROFESSOR(PROFID, PNAME, PHONE)
- ENROLLMENT(ROLLNO, COURSEID, GRADE)
- Write SQL expressions for the following queries:
- (i) Names of courses taught by 'Prof. Raju'.
- (ii) Names of students who have *not* enrolled for any course taught by 'Prof. Ganapathy'.
- (iii) For each course, name of the course and number of students enrolled for the course.
14. Assume that the relation R(P,Q,S,T,U) with FDs $P \rightarrow S$, $Q \rightarrow S$, $S \rightarrow T$, $TU \rightarrow S$, $SU \rightarrow P$ is decomposed into 5 relations: R1(P,T), R2(P,Q), R3(Q,U), R4(S,T,U) and R5(P,U). Apply the standard algorithm to test if the decomposition is a lossless-join decomposition. (9)

PART E

Answer any four full questions

15. Consider the tables R (A, B, C), T(D,E,F), S(G, H) and U(A,D, G, I) where A, D and G in U are foreign keys referring to the primary keys with the same names. Show an *initial query tree* for the following query and optimize it using the rules of heuristics: (10)

```
select B, E, G, H, I
from R, T, S, U
where R.A = U.A
and T.D = U.D and S.G = U.G
and R.C = 'TEXT' and U.I > 20 and T.E = 25
```

- 16 Consider a file with 2,00,000 records stored in a disk with fixed length blocks of size 256 bytes. Each record is of size 50 bytes. The primary key is 4 bytes and block pointer is 6 bytes. Compute the following, assuming that multi-level primary index is used as access path: (10)
- (i) Blocking factor for data records
 - (ii) Blocking factor for index records
 - (iii) Number of data blocks
 - (iv) Number of First level index blocks
 - (v) Number of levels of multi level index
17. a. Argue that two-phase locking ensures serializability. (4)
- b. Illustrate *clustering* index and *secondary* index with typical, real examples. (6)
18. a. Show the generic structure of a B+-Tree clearly indicating the types keys and pointers and their significance. (5)
- b. What is the significance of check-pointing? Illustrate with a typical example. (5)
19. a. Illustrate *lost-update* and *dirty-read* problems with suitable examples. (4)
- b. Determine if the following schedule is serializable. (6)
- r1(X), r2(Z), r1(Z), r3(X), r3(Y), w1(X), w3(Y), r2(Y), w2(Z), w2(Y)
- (Note: ri(X)/wi(X) means transaction Ti issues read/write on item X)
20. a. Write a small RDF document and show its equivalent graph structure. (4)
- b. List out any *three* salient features of Big data. (3)
- c. How is GIS databases different from conventional databases? (3)

Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
FOURTH SEMESTER B.TECH DEGREE EXAMINATION, JULY 2017

Course Code: CS208

Course Name: PRINCIPLES OF DATABASE DESIGN (CS, IT)

Max. Marks: 100

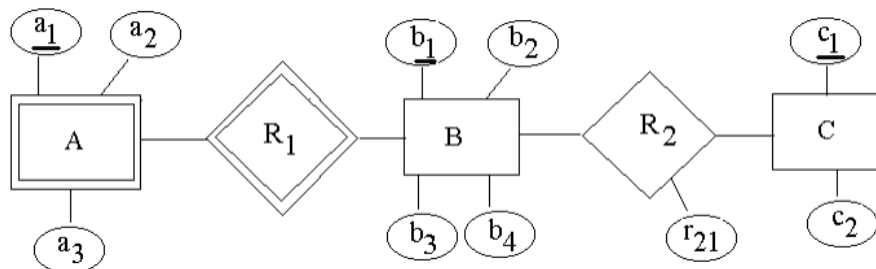
Duration: 3 Hours

Limit answers to the required points.

PART A

Answer all questions. Each carries 3 marks.

- 1 What are the responsibilities of the DBA? (3)
- 2 Define the following terms: (3)
 - i) Data model ii) Database schema iii) Meta-data
- 3 Consider the following ER diagram. Using this ER diagram create a relational database (primary keys are underlined). (3)



- 4 What are the different ways of classifying a DBMS? (3)

PART B

Answer any two questions. Each carries 9 marks.

- 5 With the help of a neat diagram explain the three-schema architecture of DBMS. (9)
- 6 Explain the following terms briefly: - (9)
 - i) Participation constraint
 - ii) Overlap constraint
 - iii) Covering constraint
- 7 Consider the following database with primary keys underlined (9)

Suppliers (sid, sname, address)

Parts (pid, pname, color)

Catalog (sid, pid, cost)

sid is the key for Suppliers, *pid* is the key for Parts, and *sid* and *pid* together form the key for Catalog. The Catalog relation lists the prices charged for parts by Suppliers.

Write relational algebra for the following queries: -

- i) Find then names of suppliers who supply some red part
- ii) Find the *sids* of suppliers who supply some red or green part
- iii) Find the *sids* of suppliers who supply some red part and some green part.

PART C*Answer all questions. Each carries 3 marks.*

- 8 What are the basic data types available for attributes in SQL? (3)
- 9 List the aggregate functions in SQL. (3)
- 10 Let $E = \{B \rightarrow A, D \rightarrow A, AB \rightarrow D\}$ is a set of Functional Dependencies. Find a minimal cover for E. (3)
- 11 Define Boyce-Codd normal form(BCNF). Give an example of a relation that is in 3NF but not in BCNF. (3)

PART D*Answer any two questions. Each carries 9 marks.*

- 12 Consider the following relations for bank database (Primary keys are underlined):

Customer (customer-name, customer-street, customer-city)

Branch (branch-name, branch-city, assets)

Account (account-number, branch-name, balance)

Depositor (customer-name, account-number)

Loan (loan-number, branch-name, amount)

Answer the following in SQL:

- i) Create tables with primary keys and foreign keys (5)
- ii) Create an assertion for the sum of all loan amounts for each branch must be less than the sum of all account balances at the branch. (4)
- 13 Given $R(A,B,C,D,E)$ with the set of FDs, $F = \{AB \rightarrow CD, ABC \rightarrow E, C \rightarrow A\}$.
- i) Find any two candidate keys of R (3)
- ii) What is the normal form of R? Justify your answer. (6)
- 14 a) What are Armstrong's axioms? (3)
- b) Write an algorithm to compute the attribute closure of a set of attributes (X) under a set of functional dependencies (F). (3)
- c) Explain three uses of attribute closure algorithm. (3)

PART E*Answer any four questions. Each carries 10 marks.*

- 15 What are the different types of single-level ordered indices? Explain. (10)
- 16 a) What is a B^+ -tree? (2)
- b) Describe the structure of both internal and leaf nodes of a B^+ -tree of order p (8)
- 17 Differentiate between static hashing and dynamic hashing. (10)
- 18 How concurrency is controlled using Timestamp Ordering algorithm. (10)
- 19 Explain the concepts behind the following: -
- i) Log-Based Recovery (5)
- ii) Deferred Database Modification. (5)
- 20 a) What are the components of GIS? (3)
- b) Explain the characteristics of data in GIS. (3)
- c) What are the constraints in GIS? (4)

Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
FOURTH SEMESTER B.TECH DEGREE EXAMINATION, APRIL 2018

Course Code: CS208

Course Name: PRINCIPLES OF DATABASE DESIGN (CS, IT)

Max. Marks: 100

Duration: 3 Hours

Limit answers to the required points.

PART A

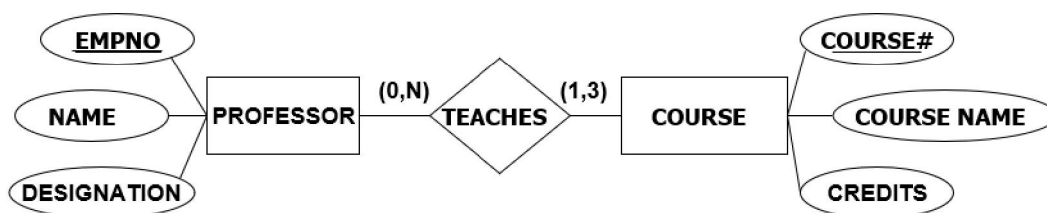
Answer all questions, each carries 3 marks.

- | | Marks |
|--|-------|
| 1 List any <i>three</i> categories of database users, highlighting any <i>one</i> important characteristic of <i>each</i> category. | (3) |
| 2 In a relationship of <i>degree 2</i> , how can we decide if an attribute of the relationship can be moved to one of the entity sets? | (3) |
| 3 Give suitable examples for <i>multi-valued</i> , <i>composite</i> and <i>multi-valued composite</i> attributes. | (3) |
| 4 Given the relations PROFESSOR(<u>PID</u> , PNAME, DEPT, SALARY) and STUDENT(<u>ROLLNO</u> , NAME, CLASS, ADVISER) where ADVISER is a foreign key to PROFESSOR, write <i>one</i> equivalent relational algebra expression for <i>each</i> of the following: | (3) |
| i. $\Pi_{PNAME}(\sigma_{SALARY > 30000}(\text{PROFESSOR}))$ | |
| ii. $\Pi_{SNAME}(\sigma_{DEPT='CSE'}(\text{STUDENT} \bowtie \text{PROFESSOR}))$
ADVISER=PID | |

PART B

Answer any two full questions, each carries 9 marks.

- 5 a) Distinguish between *total* and *partial* participation constraints with the help of real examples. (4)
- b) Describe the real-world situation described by the following ER diagram. (3)

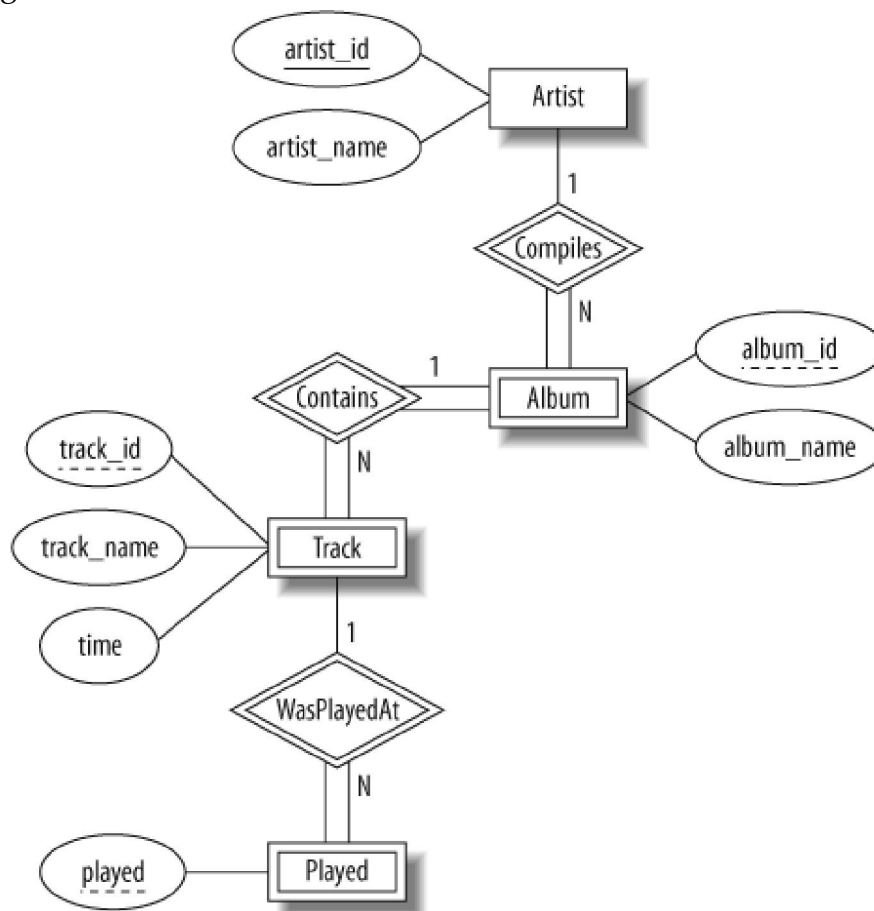


- c) Re-draw the ER diagram in 5(b) replacing the *(min,max)* notation with the conventional notation showing cardinality and participation. (2)
- 6 The relational schema for a library describing members, books and issue information is given below. Foreign keys have the same name as primary keys.
BOOKS(ACC-NO, ISBN, TITLE, EDITION, YEAR)

MEMBERS(MEMBERID, MEMBERNAME, MEMBERTYPE)
 ISSUEDTO(ACC-NO, MEMBERID, DATE OF ISSUE)

Write relational algebra expressions for the following queries:

- a) Accession Number(s) and Name(s) of third edition books published in 2018. (2)
 - b) Names and dates of issue of books taken by a member with name 'PRIYA'. (3)
 - c) Names of books *not* taken by any member. (4)
- 7 a) Use the standard synthesis procedure to generate the set of relations corresponding to the ER diagram below. Identify primary and foreign keys of the generated relations. (4)



- b) What is meant by referential integrity? How is it implemented using foreign key? Illustrate using a real example. (5)

PART C

Answer all questions, each carries 3 marks.

- 8 Consider two tables STUDENT(ROLLNO, NAME, CLASS) and ENROLLMENT(ROLLNO, COURSENAME) where ROLLNO in ENROLLMENT is a foreign key referring to STUDENT. It is required that every time a STUDENT tuple is deleted, all the ENROLLMENT tuples referring (3)

to the deleted STUDENT tuple are also deleted. Write SQL statements to specify this foreign key requirement.

- 9 Illustrate use of *assertions* with an example. (3)
- 10 When do you say that two sets of functional dependencies are *equivalent*? Give a brief explanation. (3)
- 11 What is meant by *transitive dependency*? Given an example. (3)

PART D

Answer any two full questions, each carries 9 marks.

- 12 Consider the following relations:
 FACULTY(FNO, NAME, GENDER, AGE, SALARY, DNUM)
 DEPARTMENT(DNO, DNAME, DPHONE)
 COURSE(CNO, CNAME, CREDITS, ODNO)
 TEACHING(FNO, CNO, SEMESTER)
- DNUM is a foreign key that identifies the department to which a faculty belongs. ODNO is a foreign key identifying the department that offers a course.
- Write SQL expressions for the following queries:
- a) Course numbers and names of 3-credit courses offered by 'CS' department. (2)
- b) Names of faculty members teaching *maximum* 3 courses. (3)
- c) Names of departments along with number of courses offered by each of them, in the increasing order of number of courses; exclude departments which do not offer any course. (4)
- 13 Given a relation $R(A_1, A_2, A_3, A_4, A_5)$ with functional dependencies $A_1 \rightarrow A_2 A_4$ and $A_4 \rightarrow A_5$, check if the decomposition $R_1(A_1, A_2, A_3)$, $R_2(A_1, A_4)$, $R_3(A_2, A_4, A_5)$ is lossless. (9)
- 14 a) For the relations listed below, write SQL statements for the updates that follow. (Assume suitable domains for attributes.) (5)
- ALBUMS(ALBUM-ID, ALBUM-NAME, PRODUCED-BY, YEAR)
 SONGS(SONG-ID, SONG-START, DURATION, ALBUM-ID)
- Update the year of the album with name 'SUHANA RATH' to 2018.
 Delete the album 'YADON KI BAARISH' along with all the songs in it.
- b) Briefly discuss 3NF and BCNF with suitable real examples. (4)

PART E

Answer any four full questions, each carries 10 marks.

- 15 a) Distinguish between *dense index* and *sparse index* and give examples for each. (4)

Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
FOURTH SEMESTER B.TECH DEGREE EXAMINATION, DECEMBER 2018

Course Code: CS208

Course Name: PRINCIPLES OF DATABASE DESIGN (CS, IT)

Max. Marks: 100

Duration: 3 Hours

Limit answers to the required points.

PART A

Answer all questions, each carries 3 marks.

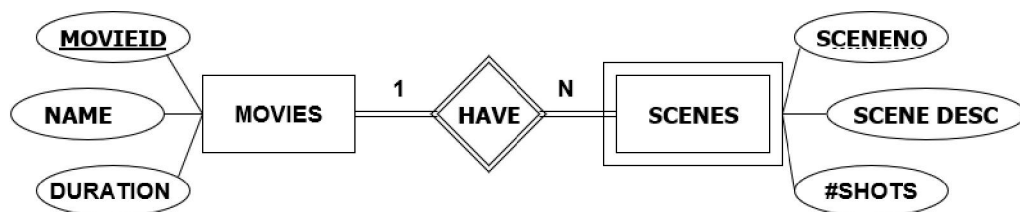
Marks

- 1 List out any *three* responsibilities of database administrators. (3)
- 2 Give *good* examples (using ER notation) for unary and ternary relationships with a *very brief* explanation. (3)
- 3 Consider a scenario where artists act in movies: an artist can act in *different* movies and movie can have *many* artists. Assuming suitable attributes show how the situation can be represented using relations with foreign keys. (A relational schema showing primary and foreign keys is sufficient. Minimal number of attributes is required.) (3)
- 4 Two *relational algebra* expressions are said to be *equivalent* if they produce exactly the same output. Consider a relation R(A,B,C,D,E) with A as its key. (3)
 - (i) What can you say about the number of tuples returned by the expression $\sigma_{A,C}(R)$?
 - (ii) Write *two* relational algebra expressions *equivalent* to $\sigma_{A<20}(\sigma_{D>30}(R))$.

PART B

Answer any two full questions, each carries 9 marks.

- 5 a) Briefly explain the concepts of *physical data independence* and *logical data independence* with a typical real-world example for each. (5)
- b) In the following ER diagram, how can we replace the entity set SCENE with an attribute of the entity set MOVIE? Draw the new ER diagram. (4)



- 6 The relational database schema below represents certain information about albums, songs in the albums and singers of those songs. Foreign keys are given the *same* name as primary keys for easy identification. (9)

ALBUMS(ALBUM#, ALBUM-NAME, PRODUCED-BY, YEAR)

SONGS(SONG#, SONG-START, DURATION, ALBUM#)

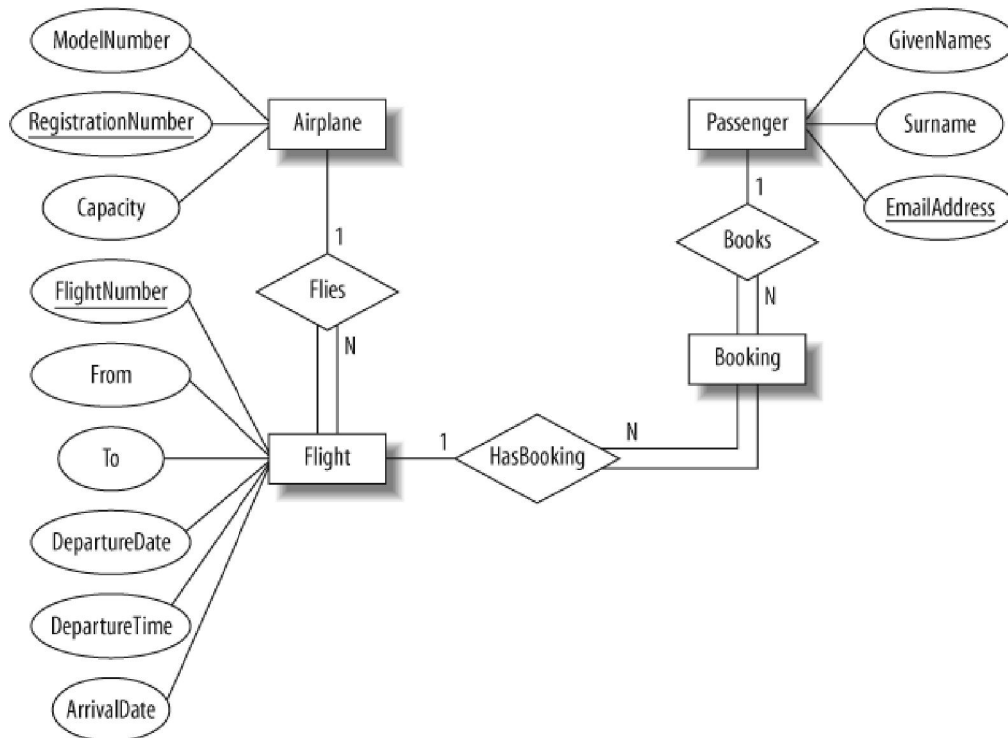
SUNGBY(ARTISTNAME, SONG#)

In the context of the schema, write relational algebra expressions for the following queries:

- (a) Names of albums produced by 'HMV' in the year 2018. (b) Names of albums in which an artist with name, 'AVANTHIKA' sung. (c) Names of albums in which

all the artists have sung songs.

- 7 a) Use the standard synthesis procedure to generate the set of relations corresponding to the ER diagram below. Identify primary and foreign keys of the relations (4)



- b) In the relational schema for a library given below, foreign keys have the same name as primary keys. Draw an ER diagram for the schema, clearly marking keys and cardinality constraints. (5)

BOOKS(ACC-NO, TITLE, EDITION, YEAR)

MEMBERS(MEMBERID, MEMBERNAME, MEMBERTYPE)

ISSUEDTO(ACC-NO,MEMBERID, DATEOFISSUE)

PART C

Answer all questions, each carries 3 marks.

- 8 Illustrate DELETE and UPDATE clauses using typical examples. (3)
- 9 Given two tables STUDENT(ROLLNO,NAME,CLASS) and ENROLLMENT(ROLLNO,COURSENAME) where ROLLNO in ENROLLMENT refers to STUDENT, what does the following SQL statement return? (3)
 SELECT COURSENAME FROM ENROLLMENT WHERE ROLLNO = ALL (SELECT ROLLNO FROM STUDENT)
- 10 Define super key and minimal super key and illustrate using good examples. (3)
- 11 Given a relation R(A,B,C,D,E,F) with functional dependencies $A \rightarrow B$, $B \rightarrow D$, $D \rightarrow EF$, $F \rightarrow A$, compute $\{D\}^+$ and $\{EF\}^+$. (3)

PART D

Answer any two full questions, each carries 9 marks.

- 12 Consider the following relations: (9)

FACULTY(FNO, NAME, GENDER, AGE, SALARY, DNUM)
 DEPARTMENT(DNO, DNAME, DPHONE)
 COURSE(CNO, CNAME, CREDITS, ODNO)
 TEACHING(FNO, CNO, SEMESTER)

DNUM is a foreign key that identifies the department to which a faculty belongs. ODNO is a foreign key identifying the department that offers a course. Write SQL expressions for the following queries:

- (a) Names and department names of faculty members. (b) Names of faculty members not offering any course. (c) Names of departments offering more than three courses, in alphabetic order.
- 13 Given a relation R(A,B,C,D,E,F,G, H) with keys BD and C and functional dependencies $D \rightarrow G$, $E \rightarrow F$ and $H \rightarrow C$, decompose the R into the highest normal form possible. (9)

- 14 a) For the relations listed below, write SQL statements to create the database schema. (5)

Assume suitable data types.

ALBUMS(ALBUM#, ALBUM-NAME, PRODUCED-BY, YEAR)

SONGS(SONG#, SONG-START, DURATION, ALBUM#)

SUNGBY(ARITISTNAME, SONG#)

- b) Why the following is table, TRAININFO, not in 1NF? How can we make it 1NF? (4)

<u>Route No</u>	Rout Name	Distance	Trains	
			Train No	Name
5	MAS-TVC	179	12475	AP Express
			13457	Tvm Mail
9	TVC-DELHI	2781	16345	ND Express
			12461	Kerala Express
			16277	NZM Mail

PART E

Answer any four full questions, each carries 10 marks.

- 15 a) Define the following: (i) physical record (ii) logical record (iii) blocking factor. (3)

- b) There are 12000 records in a data file. Each record in the file is of 75 bytes. (7)

Compute the number of block accesses if (i) Single level secondary index is available on a field of size 15 bytes. (ii) Multilevel index is available on the same field.

Assume that the block size is 394 bytes, that un-spanned organization is used and that block and record pointers are 5 and 7 bytes, respectively.

- 16 a) How is *clustering index* different from *primary index*? (3)

- b) Illustrate structure of B-Tree and B+-Tree and explain how they are different. (5)

- c) Give *anon-canonical query tree* for the expression $\square_{A,B}(\square_{D<20}(R*S))$ on the relations R(A,B,C,D) and S (D,E). (2)
- 17 a) Consider the following tables representing enrolment of students to courses: (10)
STUDENT (ROLLNO, NAME, AGE, ADDRESS, EMAIL), COURSE(CNO, CNAME, AREA) and ENROLEMENT(ROLLNO, CNO,GRADE), where ROLLNO and CNO in ENROLMENT are foreign keys referring to the primary keys with the same names.
Show an initial query tree for the following query and optimize it using the rules of heuristics. Assume that CNAME is a candidate key of COURSE.
SELECT CNAME, NAME, EMAIL, GRADEFROM STUDENT, COURSE, ENROLMENTWHERE COURSE.CNO=ENROLMENT.CNO AND STUDENT.ROLLNO = ENROLLMENT.ROLLNO AND COURSE.CNAME = 'dbs'
- 18 a) Check if the following schedules are conflict-serializable using precedence graph. (4)
If so, give the equivalent serial schedule(s). $r3(X), r2(X), w3(X), r1(X), w1(X)$.
(Note: $ri(X)/wi(X)$ means transaction T_i issues read/write on item X.)
- b) Discuss *dirty-read* and *lost-update* problems with the help of examples. (6)
- 19 a) How is *strict two-phaselocking* different from *standard two-phase locking* protocol? (2)
- b) With the help of suitable logs, show how recovery is done in *deferred database modification scheme*. (5)
- c) What is the significance of *check-pointing*? (3)
- 20 a) Give a simple RDF document and show the corresponding graph structure. (3)
- b) Write an explanatory note on Big Data. (4)
- c) Highlight any *three* salient features of GIS. (3)

Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
FOURTH SEMESTER B.TECH DEGREE EXAMINATION, MAY 2019

Course Code: CS208

Course Name: PRINCIPLES OF DATABASE DESIGN (CS, IT)

Max. Marks: 100

Duration: 3 Hours

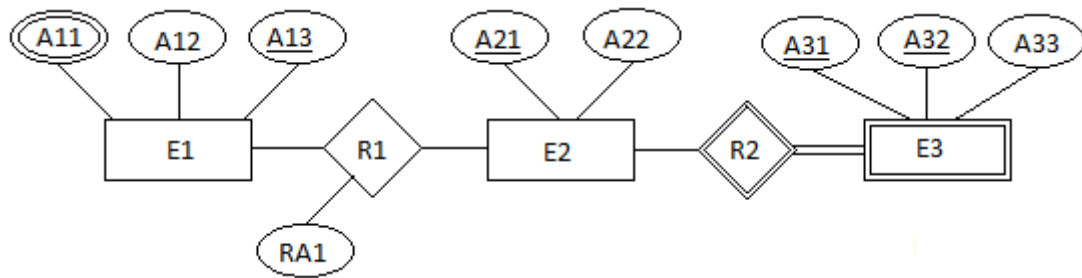
Limit answers to the required points.

PART A

Answer all questions, each carries 3 marks.

Marks

- 1 Illustrate with an example, the difference between the conceptual data models and the physical data models. (3)
- 2 How is weak entity type different from a strong entity type? Give an example. (3)
- 3 What is entity integrity constraint? Why is it important? (3)
- 4 Using the following ER diagram, create a relation database. Give your assumptions. (3)



PART B

Answer any two full questions, each carries 9 marks.

- 5 a) With the help of an example, compare DML and DDL. (4)
- b) What are logical data independence and physical data independence? What is the difference between them? Which of these harder to realize? Why? (5)
- 6 Design an ER diagram to represent the following scenario: (9)
 A company has many employees working on a project. An employee can be part of one or more projects. Each employee works on a project for certain amount of time. Assume suitable attributes for entities and relations. Mark the primary key(s) and the cardinality ratio of the relations.
- 7 Consider the following relations for a database that keeps track of business trips of salespersons in a sales office:
 SALESPERSON(Ssn, Name, StartYear, DeptNo)

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

EXPENSE(TripId, AccountNo, Amount)

- a) A trip can be charged to one or more accounts. Specify the foreign keys for this schema, stating any assumptions you make. (3)
- b) Write relation algebra expression to get the details of salespersons who have travelled between Mumbai and Delhi and the travel expense is greater than Rs. 50000. (3)
- c) Write relation algebra expression to get the details of salesperson who had incurred the greatest travel expenses among all travels made. (3)

PART C

Answer all questions, each carries 3 marks.

- 8 With the help of an example, illustrate the use of SQL TRIGGER. (3)
- 9 List the basic data types available for defining attributes in SQL? (3)
- 10 Consider a relation $R=\{A,B,C,D,E,F\}$ and a set of functional dependencies $F=\{A \rightarrow BC, C \rightarrow BD, BF \rightarrow E, F \rightarrow D\}$. Find the closure of A. Is A a candidate key? Justify. (3)
- 11 What are fully functional dependencies and partial functional dependencies? Give an example to distinguish between these? (3)

PART D

Answer any two full questions, each carries 9 marks.

- 12 a) Consider the following table MARKS. Why is the table not in 1NF? Reconstruct the table so that it is in 1NF. (5)

Roll No.	Name	Marks	Subject	
			Code	Name
1001	Tom	42	M001	Maths
		34	C002	Chemistry
		37	P003	Physics
1057	Sam	21	M001	Maths
		25	C002	Chemistry
		34	P003	Physics
1001	Tom	45	M001	Maths
		48	C002	Chemistry
		44	P003	Physics

- b) When does a relational scheme is said to be in 3NF? How is BCNF different from 3NF? (4)
- 13 a) List aggregate functions of SQL. (3)

- b) Given a relation $R(A,B,C)$. Find the minimal cover of the set of functional dependencies given; (3)
- $F = \{A \rightarrow BC, B \rightarrow C, A \rightarrow B, AB \rightarrow C\}$
- c) What is the lossless (or nonadditive) join property of decomposition? Why is it important? (3)
- 14 Consider the relation $R = \{A, B, C, D, E, F, G, H\}$ and the set of functional dependencies $F = \{A \rightarrow DE, B \rightarrow F, AB \rightarrow C, C \rightarrow GH, G \rightarrow H\}$. What is the key for R? Decompose R into 2NF and then 3NF relations. (9)

PART E

Answer any four full questions, each carries 10 marks.

- 15 a) Suppose that we have an ordered file with 400,000 records stored on a disk with block size 4,096 bytes. File records are of fixed size and are unspanned, with record length 200 bytes. How many blocks are needed for the file? Approximately, how many block accesses are required for a binary search in this file? On an average, how many block accesses are required for a linear search, if the file is nonordered? (6)
- b) Based on question 15.a, give an example to illustrate that indexing can improve the search time. (4)
- 16 a) Explain the structure of an internal node and a leaf node in a B+-tree. (5)
- b) Illustrate with an example how searching for a record with search key field value is done using a B+-Tree. (5)
- 17 Why Concurrency Control Is Needed? What are the different types of problems we may encounter when two transactions run concurrently? Illustrate each problem with suitable examples. (10)
- 18 a) What are the desirable properties of transactions? Explain. (4)
- b) "If every transaction in a schedule follows the two-phase locking protocol, the schedule is guaranteed to be serializable", justify the statement. (3)
- c) What are the different types of lock that are commonly used in concurrency control? (3)
- 19 a) Consider the following tables representing courses taken by instructors in an institute: (10)
- INSTRUCTOR(ID, NAME, DEPT, SALARY)
- TEACHES(ID, COURSE-ID, SEMESTER, YEAR)
- COURSE(COURSE-ID, TITLE, DEPT, CREDITS)
- where, ID and COURSE-ID are foreign keys referring to the primary keys with the

same names. Show an initial query tree for the following query and optimize it using the rules of heuristics. Assume that TITLE is a candidate key of COURSE.

```
SELECT NAME, TITLE, SEMESTERYEAR
FROM INSTRUCTOR, COURSE, TEACHES
WHERE COURSE.COURSE-ID=TEACHES.COURSE-ID AND
TEACHES.ID = INSTRUCTOR.ID AND INSTRUCTOR.DEPT = 'MATHS'
AND TEACHES.DEPT = INSTRUCTOR.DEPT
```

- 20 a) Write a short note on Big Data. (4)
- b) What is a semantic web technology? How is it relevance? (3)
- c) How does RDF support semantic web technology? (3)

Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY
FOURTH SEMESTER B.TECH DEGREE EXAMINATION(S), DECEMBER 2019

Course Code: CS208

Course Name: PRINCIPLES OF DATABASE DESIGN

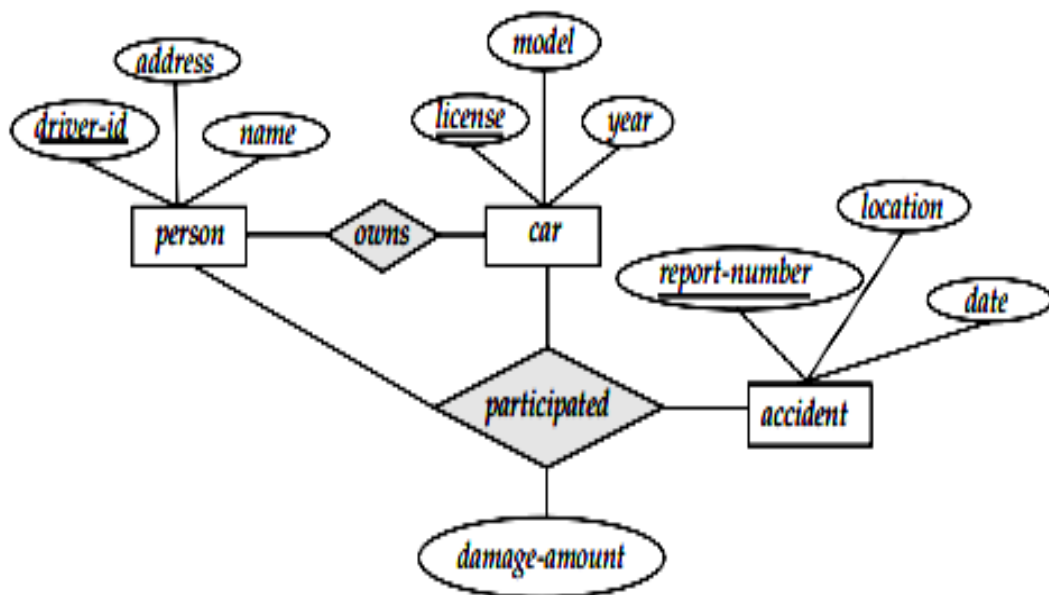
Max. Marks: 100

Duration: 3 Hours

PART A

Answer all questions, each carries 3 marks.

- | | | Marks |
|---|--|-------|
| 1 | List three significant differences between a file-processing system and a DBMS. | (3) |
| 2 | When is the concept of weak entity used in data modelling? Define the following terms:
(i) Identifying relationship type
(ii) Owner entity type | (3) |
| 3 | Consider the relational database shown below:
employee (<u>person-name</u> , street, city)
works (<u>person-name</u> , <u>company-name</u> , salary)
company (<u>company-name</u> , city)
manages (<u>person-name</u> , <u>manager-name</u>)
Define the integrity constraints in the given relation. State the assumptions used. | (3) |
| 4 | Design a relational database corresponding to the following E-R diagram. | (3) |



PART B

Answer any two full questions, each carries 9 marks.

- | | | |
|---|--|-----|
| 5 | a) Explain the characteristics of the database approach. | (3) |
| | b) Explain three schema architecture with figure. | (6) |
| 6 | a) Who all are the actors on the scene in DBMS and explain their roles? | (3) |
| | b) Explain the left outer join, right outer join and full outer join operations with examples. | (6) |

- 7 Consider the following schema: (9)
 Suppliers(sid, sname, address)
 Parts(pid, pname, colour)
 Catalog (sid, pid, cost)
 The key fields are underlined. Foreign key in Catalog are sid and pid referring to Supplier and Parts, respectively. Write the relational algebra expression for the following queries:
 (i) Find the sids of suppliers who supply some red or green part.
 (ii) Find the sids of suppliers who supply every part.
 (iii) Find the names of suppliers who supply some red part.

PART C*Answer all questions, each carries 3 marks.*

- 8 Illustrate the concept of trigger in SQL with an example. (3)
 9 Give any three examples to illustrate the schema modification statements in SQL. (3)
 10 Explain any three uses of attribute closure algorithm. (3)
 11 Consider the following set F of functional dependencies for relation schema (3)
 $R = (A, B, C, D, E)$.
 $F = \{A \rightarrow BC, CD \rightarrow E, B \rightarrow D, E \rightarrow A\}$
 Compute the canonical cover of F.

PART D*Answer any two full questions, each carries 9 marks.*

- 12 a) What is an assertion? (2)
 b) Write an assertion for the bank database to ensure that the assets value for the Perryridge branch is equal to the sum of all the amounts lent by the Perryridge branch. (7)
 The schema for branch and loan are shown below:
 branch(branch_name, branch_city, assets)
 loan(loan_number, branch_name, amount)
 13 a) Explain the difference between BCNF and 3NF with an example. (4)
 b) Given below are two sets of FDs for a relation R(A,B,C,D,E). Are they equivalent? (5)
 $F_1 = \{A \rightarrow B, AB \rightarrow C, D \rightarrow AC, D \rightarrow E\}$
 $F_2 = \{A \rightarrow BC, D \rightarrow AE\}$
 14 a) Give suitable example for nested sub queries. Write the advantages of nested query? (4)
 b) Suppose that we decompose the schema $R = (A, B, C, D, E)$ into (5)
 $R_1(A, B, C)$
 $R_2(A, D, E)$
 Test whether the given decomposition is a lossless-join decomposition, if the following set F of functional dependencies holds in R:
 $F = \{A \rightarrow BC, D \rightarrow E, B \rightarrow D, E \rightarrow A\}$

PART E*Answer any four full questions, each carries 10 marks.*

- 15 a) What is the main difference between a primary index and a clustering index? Give examples. (3)
 b) How does multilevel indexing improve the efficiency of searching an index file? Explain your answer. (4)
 c) Is it possible in general to have two primary indices on the same relation for different search keys? Explain your answer. (3)
 16 a) Draw the structure of a B+-tree with q-1 search values. Differentiate between internal nodes and leaf nodes of B+-tree. (5)
 b) Construct a B+-tree for the following set of key values: (5)
 (2, 3, 5, 7, 11, 17, 19, 23, 29, 31)

Assume that the tree is initially empty and values are added in ascending order. Construct B+-trees for the case where the number of pointers that will fit in one node is four.

- 17 a) What is Big Data? (3)
b) What is the significance of Big Data in current data management scenarios? (3)
c) How is Big Data different from traditional data sources? (4)
- 18 a) What are the different types of locks used in concurrency control? (2)
b) How conversions of locks are achieved in concurrency control? (3)
c) Explain how we can guarantee serializability by two phase locking protocol. (5)
- 19 a) Consider the following relations: (10)
Instructor(IID, Iname,Iage,Idept,Isalary)
Teaches(IID,CID, semester)
Course(CID,Ctitle,credit)
The foreign keys in Teaches are same as the primary key names. Perform heuristics-based query optimization on the following SQL query.
SELECT IID
FROM Instructor, Teaches, Course
WHERE Instructor.IID = Teacher.IID AND Course.CID = Teaches.CID AND
Teaches.semester = "Even" AND Course.title = "DBMS";
- 20 a) Explain the components of GIS systems. (4)
b) What are the characteristics of data in GIS. (6)

Reg No.: _____

Name: _____

APJ ABDUL KALAM TECHNOLOGICAL UNIVERSITY

Fourth semester B.Tech examinations (S), September 2020

Course Code: CS208**Course Name: PRINCIPLES OF DATABASE DESIGN (CS, IT)**

Max. Marks: 100

Duration: 3 Hours

Limit answers to the required points.**PART A*****Answer all questions, each carries 3 marks.***

Marks

- | | | |
|---|---|-----|
| 1 | Illustrate three schema architecture with a suitable diagram. | (3) |
| 2 | With a help of an example, elaborate weak entity set and strong entity set. | (3) |
| 3 | Differentiate DDL and DML with suitable example. | (3) |
| 4 | With suitable example, define integrity constraint? | (3) |

PART B***Answer any two full questions, each carries 9 marks.***

- | | | |
|---|--|-----|
| 5 | Design an ER diagram for the given scenario; | (9) |
|---|--|-----|

Suppose that you are designing a schema to record information about reality shows on TV. Your database needs to record the following information:

_ For each reality show, its name, genre, basic_info and participants name. Any reality show has at least two or more participants.

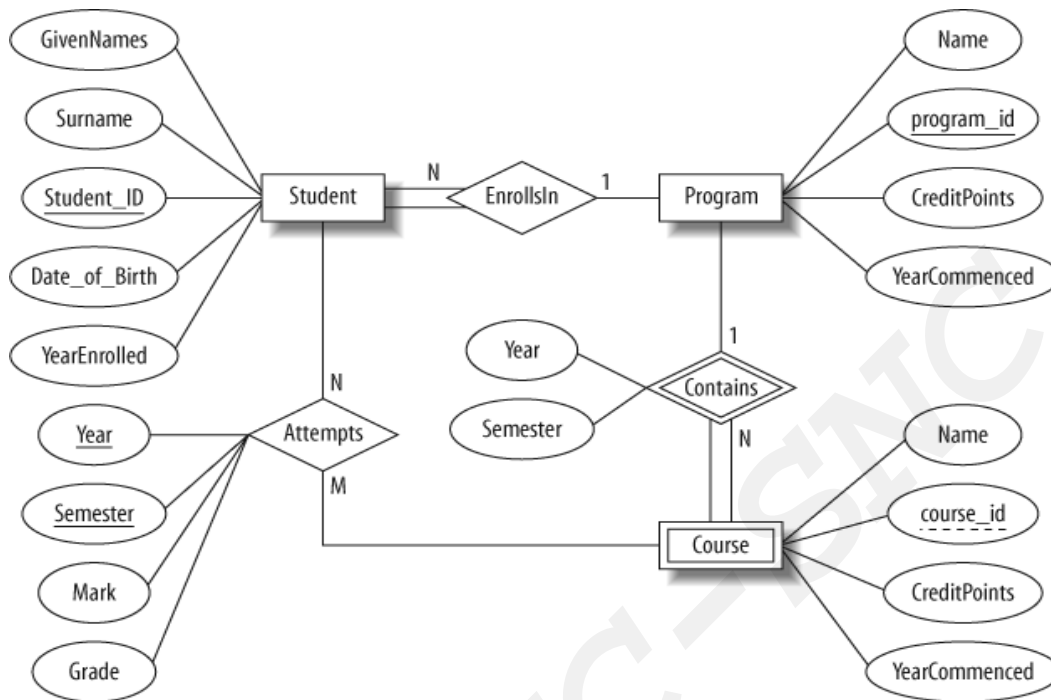
- For each producer, the company name, company country. A show is produced by exactly one producer. And one producer produces exactly one show.

- For each television, its name, start year, head office. A television may broadcasts multiple shows. Each show is broadcasted by exactly one television.

-For each user, his/her username, password, and age. A user may rate multiple shows, and a show may be rated by multiple users. Each rating has a score of 0 to 10.

6 Covert the following ER Model to Relational Model

(9)



7 a) Differentiate Natural join, Equi-join and Left outer join

(3)

7 b) Consider the schema given below.

(6)

employee (person-name, street, city)

works (person-name, company-name, salary)

company (company-name, city)

manages (person-name, manager-name)

Write relational algebra queries for the following questions

a. Find the names and cities of residence of all employees who work for First Bank Corporation.

b. Find the names, street address, and cities of residence of all employees who work for First Bank Corporation and earn more than \$10,000 per annum.

c. Find the names of all employees in this database who live in the same city as the company for which they work.

PART C*Answer all questions, each carries 3 marks.*

8 With suitable example, list aggregate functions in sql

(3)

9 What is the importance of views in sql? Explain with suitable example.

(3)

10 Illustrate different anomalies in designing a database.

(3)

11 How can we conclude two FDs are equivalent?

(3)

PART D

Answer any two full questions, each carries 9 marks.

- 12 a) Consider the schema given below. (9)
- person (driver-id, name, address)
car (reg-no, model, year, driver-id)
accident (report-number, date, location)
participated (driver-id, reg-no, report-number, damage-amount)
- Write SQL queries for the following
- a. Find the name of driver, who is drives the car with reg-no='AABB2000'. Find the total number of people who were involved in car accidents in 01-01-1989.
- c. Find the number of accidents in which the cars belonging to “John Smith” were involved.
- d. Update the damage amount for the car with reg-no “AABB2000” in the accident with report number “AR2197” to \$3000.
- 13 a) Compute the closure of the following set F of functional dependencies for relation schema R = (A, B, C, D, E). (3)
- $A \rightarrow BC$
 $CD \rightarrow E$
 $B \rightarrow D$
 $E \rightarrow A$
- List the candidate keys for R.
- b) Define 3NF and BCNF. Let R (A, B, C, D, E) be a relational schema in which the following functional dependencies are known to hold: $AB \rightarrow C$, $C \rightarrow E$ and $E \rightarrow D$. Identify the highest normal form (6)
- 14 a) Illustrate triggers with a suitable example. (4)
- b) R = (A, B, C, D, E). We decompose it into R1 = (A, B, C), R2 = (C, D, E). The set of functional dependencies is: $A \rightarrow BC$, $CD \rightarrow E$, $B \rightarrow D$, $E \rightarrow A$. Check whether this decomposition is a lossless join decomposition or not. (5)

PART E

Answer any four full questions, each carries 10 marks.

- 15 a) Compare primary indexing, secondary indexing and clustered indexing with suitable diagram. (6)
- b) Define the structure of B+ tree (4)

- 16 a) Database system implementers have paid much more attention to the ACID properties than have file-system implementers. Why might this be the case? (6)
- b) Explain deferred database modification with an example. (4)
- 17 a) Consider a file with 450000 records . Each record size is 125 bytes and block is 1000 bytes. The primary key of the file is 10 bytes and record pointer size is 6 bytes. (10)
- 1) Calculate number of index block required in case of primary indexing
- 2) Calculate number of index blocks required in case of multilevel indexing.
- 18 a) Differentiate serial and concurrent schedules. Elaborate Conflict serializability with suitable example. (6)
- b) Illustrate two phase locking (4)
- 19 a) Why recovery is needed in transaction processing (5)
- b) Illustrate heuristic query optimization with suitable example. (5)
- 20 a) Write short notes on (10)
- 1) Big Data
- 2) Biological Database
- 3) GIS
