IV

Databases

Contents

Description Sheet

DATABASES

Chapter-1 : Introduction to DBMS and Integrity Constraints and ER Model

- What is DBMS?
- Limitation of Flat Files
- Relation Scheme
- Relation Instance
- Candidate Key
- Super Key
- Primary and Alternate Key
- Foreign Key
- Participation
- Cardinalities
- Minimization of ER Diagram

Chapter-2 : Normalization

- Functional Dependencies
- Attribute Closure
- Amstrong Properties
- Membership Test
- Testing of Equality
- Properties of Decomposition
- First Normal Form
- Second Normal Form
- Third Normal Form
- BCNF Normal Form
- Decomposition of Relation
- Multivalue Function Dependency
- Fourth Normal Form
- Minimal Cover or Canonical Cover

Chapter-3 : Relational Algebra, Tuple Relational Calculus and SQL

- Relation Algebra
- SQL
- TRC

Chapter-4 : Transaction and Concurrency Control

- Transaction
- Concurrency Control
- ACID Property
- Serializability
- Problems because of Concurrent Execution
- Classification of Scheduling
- Locking Protocol
- 2PL
- Time Stamp Ordering Protocol
- Thomas Write Time Stamp Protocol

Chapter-5 : File Structure and Indexing

- Types of Record
- Record Organization
- Categories of Index
- Types of Index
- B^+ Tree
- B Tree

■■■■

1 Introduction to DBMS and Integrity Constraints and ER Model

Multiple Choice Questions

Q.1 Consider a relation ***geq*** which represents "greater than or equal to", that is $(x, y) \in geq$ only if $y \geq x$:

(CREATE TABLE *geq*

lb integer NOT NULL,

ub integer NOT NULL,

primary key *lb*

Foreign key (*ub*) references *geq* on delete cascade);

Which of the following is possible if a tuple (x, y) is deleted?

(a) A tuple (z, w) with $z > y$ is deleted

(b) A tuple (z, w) with $z > x$ is deleted

(c) A tuple (z, w) with $w < x$ is deleted

(d) The deletion of (x, y) is prohibited

Q.2 The following table has two attributes A and C where A is the primary key and C is the foreign key referencing A with on delete cascade.

A	2	3	4	5	7	9	6
C	4	4	3	2	2	5	4

The set of all tuples that must be additionally deleted to preserve referential integrity when the tuple (2, 4) is deleted is

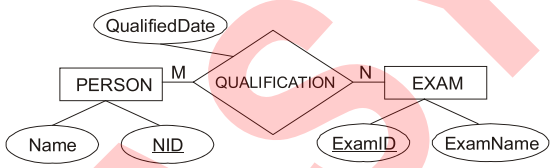
(a) (3, 4) and (6, 4)

(b) (5, 2) and (7, 2)

(c) (5, 2), (7, 2) and (9, 5)

(d) (3, 4), (4, 3) and (6, 4)

Q.3 Consider the following Entity Relationship Diagram (ERD).



Which of the following possible relations will not hold if the above ERD is mapped into a relation model?

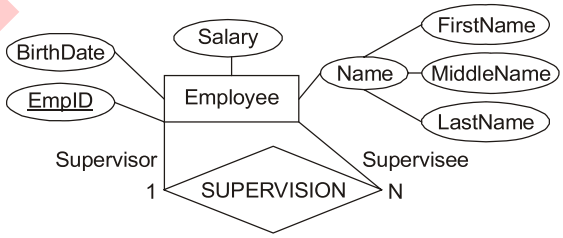
(a) Person (NID, Name)

(b) Qualification (NID, ExamID, QualifiedDate)

(c) Exam (ExamID, NID, ExamName)

(d) Exam (ExamID, ExamName)

Q.4 Consider the following ERD diagram depicting the relationship of an employee and supervisor.



Which of the possible relations if the above ERD is mapped into a relational model?

(a) Employee (EmpID, BirthDate, Salary, Name(FirstName, MiddleName, LastName))

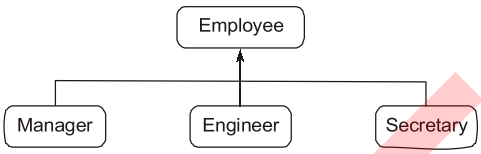
(b) Supervision (EmpID, BithDate, Salary, Name(FirstName, MiddleName, LastName), EmpId)

(c) Supervisor (SupervisorID, BirthDate, Salary, Name(FirstName, MiddleName, LastName), EmpID), {EmpID})

(d) Employee (EmpID, BirthDate, Salary, Name(FirstName, MiddleName, LastName), SupervisorID)

Q.5 It is desired to design an object-oriented employee record system for a company. Each employee has a name, unique id and salary. Employees belong to different categories and their salary is determined by their category. The functions getName, getId and compute Salary are required. Given the class hierarchy below, possible locations for these functions are:

1. getId is implemented in the superclass
2. getId is implemented in the subclass
3. getName is an abstract function in the superclass
4. getName is implemented in the superclass
5. getName is implemented in the subclass
6. getSalary is an abstract function in the superclass
7. getSalary is implemented in the superclass
8. getSalary is implemented in the subclass



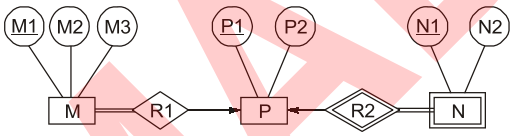
Choose the best design

(a) 1, 4, 6 and 8 (b) 1, 4 and 7
(c) 1, 3, 5, 6 and 8 (d) 2, 5 and 8

[GATE-2004]

Common Data for Q.6 & Q.7

Consider the following ER diagram:



Q.6 The minimum number of table needed to represent M, N, P, R1, R2 is

(a) 2 (b) 3
(c) 4 (d) 5

[GATE-2008]

Q.7 Which of the following is a correct attribute set for one of the tables for the correct answer to the above question?

(a) {M1,M2,M3,P1} (b) {M1, P1, N1, N2}
(c) {M1, P1, N1} (d) {M1, P1}

Q.8 Consider the following Entity Relationship Diagram (ERD), where two entities E1 and E2 have a relation R of cardinality 1 : m.

The attributes of E1 are A11, A12 and A 13 where A11 is key attribute. The attributes of E2 are A21, A22, and A23 where A21 is the key attribute and A23 is a multi-valued attribute. Relation R does not have any attribute. A relational database containing minimum number of tables with each table satisfying the requirements of the third normal form (3NF) is designed from the above ERD. The number of tables in the database is

(a) 2 (b) 3
(c) 5 (d) 4

[GATE-2004]

Q.9 The relationship between two entity types A and B is 1:1 and the relationship is optional at the A end. Only 50% of B entities are related to an A entity. Now consider mapping these entity types into relations.

Select the best statement from the following list.

(a) A and B should be kept separate with the foreign key in the A relation.
(b) A and B should be kept separate with a foreign key in both A and B.
(c) A and B should be kept separate with the foreign key in the B relation.
(d) None of the above

Q.10 1 : N relationship in E-R diagram is implemented in relation model as

(a) foreign keys are added on both sides
(b) relation corresponding to '1' side is modified to include foreign key of the relation on the N side
(c) primary keys are added on both sides
(d) relation corresponding to 'N' side is modified to include foreign key of the relation on the '1' side.

Q.11 A relation (from the relational database model) consists of a set of tuples, which implies that

(a) relational model supports multi-valued attributes whose values can be represented in sets.

(b) for any two tuples, the values associated with all of their attributes may be the same.

(c) for any two tuples, the value associated with one or more of their attributes must differ.

(d) all tuples in a particular relation may have different attributes.

Q.12 Consider the following statements.

1. An entity integrity constraint states that no primary key value can be null.
2. A referential integrity constraint is specified between two relations.
3. A foreign key cannot be used to refer to its own relation.

Identify which of the above statements is/are correct?

(a) Only 1 (b) Only 2

(c) Only 2 and 3 (d) Only 1 and 2

Q.13 Given the following statements:

S_1: A foreign key declaration can always be replaced by an equivalent check assertion in SQL.

S_2: Given the table R(*a*, *b*, *c*) where *a* and *b* together form the primary key, the following is a valid table definition.

CREATE TABLE S (
a INTEGER,
d INTEGER,
e INTEGER,
PRIMARY KEY (*d*),
FOREIGN KEY (*a*) references R)

Which one of the following statements is **CORRECT**?

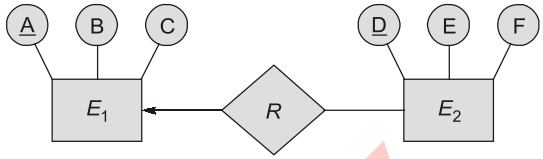
(a) S_1 is TRUE and S_2 is FALSE

(b) Both S_1 and S_2 are TRUE

(c) S_1 is FALSE and S_2 is TRUE

(d) Both S_1 and S_2 are FALSE

Q.14 Consider the following ER diagram:

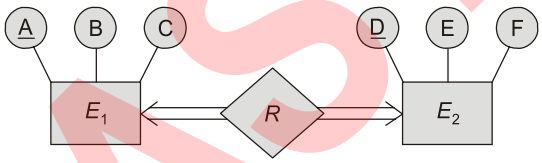


How many minimum RDBMS tables which satisfy 1NF?

(a) 1 (b) 2

(c) 3 (d) 4

Q.15

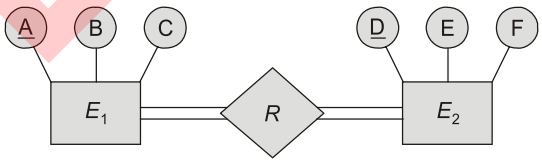


How many minimum relations required which satisfy 1NF?

(a) 1 (b) 2

(c) 3 (d) 4

Q.16



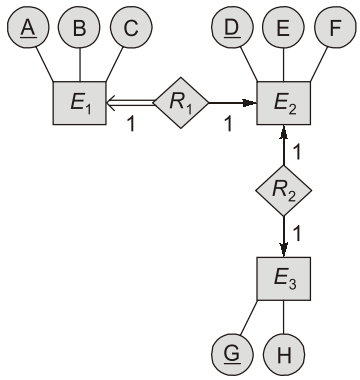
How many minimum relations required which satisfy 2NF?

(a) 1 (b) 2

(c) 3 (d) 4

Common Data for Q.17 & Q.18

Consider the following ER diagram:



Q.17 How many foreign keys required for the minimized ER diagram into relations?

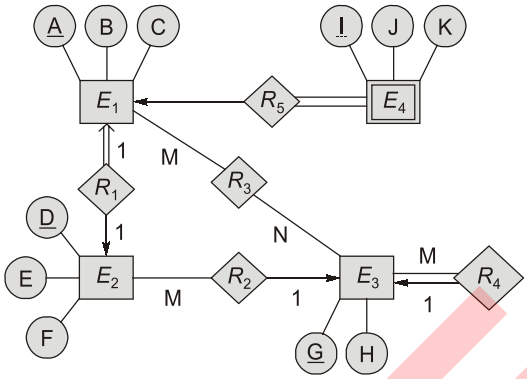
(a) 2 (b) 3
(c) 4 (d) 1

Q.18 How many total attributes required for the minimized ER diagram into relations?

(a) 6 (b) 8
(c) 9 (d) 10

Common Data for Q.19 & Q.20

Consider the following ER diagram:



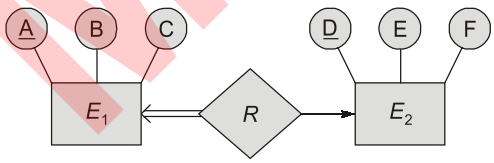
Q.19 How many minimum relations required for the above ER diagram?

(a) 3 (b) 4
(c) 5 (d) 6

Q.20 How many minimum foreign keys required for minimized ER diagram into relations?

(a) 3 (b) 4
(c) 5 (d) 5

Q.21 Which of the following is minimized relation key constraint of the of below ER diagram.



(a) A key with no Null's
D key with no Null's

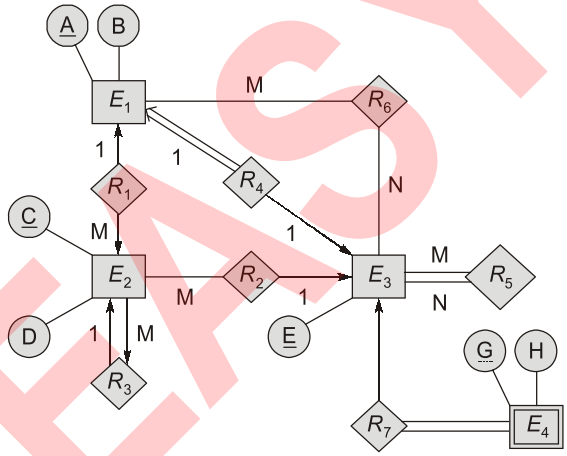
(b) A key with allowed Null's
D key with no Null's

(c) A key with no Null's
D key with allowed Null's

(d) A key with allowed Null's
D key with allowed Null's

Common Data for Q.22, 23 & Q.24

Consider the following ER diagram:



Q.22 How many minimum relations required for the above ER diagram?

(a) 5 (b) 4
(c) 6 (d) 7

Q.23 How many foreign keys required for the minimized relations of the above ER diagram?

(a) 7 (b) 8
(c) 9 (d) 5

Q.24 How many total number of attributes for minimized relations of the above ER diagram

(a) 15 (b) 14
(c) 16 (d) 17

Q.25 Consider the following relation:
$R\{A_1, A_2, ..., A_n\}$ n attributes $\{A_1, A_2, ..., A_m\}$ are m attributes of relation R are m simple candidate keys $[m \le n]$.
How many super keys possible in relation R?

(a) $(2^n - 1)m$ (b) $2^m \cdot 2^{n-m}$
(c) $2^m(2^{n-m} - 1)$ (d) $(2^{m-1}) \cdot 2^{n-m}$

Q.26 Which of the following statement true about view's?

(a) Views allowed to update if defined over single relation without aggregation.

(b) Views allowed to update if defined over multiple relations without aggregation.

(c) Views allowed to update if defined over single relation even with aggregation.

(d) View's allowed to update if defined over multiple relations with aggregation.

Q.27 Which of the statement true for logical data independence?

(a) Changes of physical file may not effect table schemas.

(b) Changes of conceptual definition of table may not effect application programming.

(c) Changes of storage file index may pot effect table definition (conceptual schema).

(d) None of the above

Q.28 Which of the following statement true for physical data independence?

(a) Changes of conceptual definition of the table may not effect application program.

(b) Changes of physical storage file may not effect conceptual definition of table.

(c) Changes of view definition not effect conceptual definition of the table.

(d) None of the above

Q.29 Which of the following statement true for external schema of levels of abstraction?

(a) Allows data access to be customized at the level of individual users.

(b) Describes all data that is actually stored in the database.

(c) Summarize how the relation described in the conceptual schema are actual on disk.

(d) All the above

Q.30 Which of the following statement true for conceptual (logical) schema of levels of abstraction?

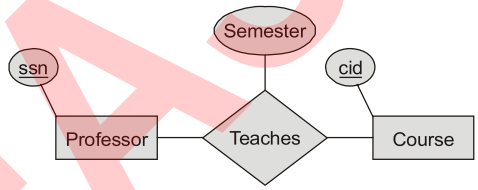
(a) Allows data access to be customized at the level of individual users.

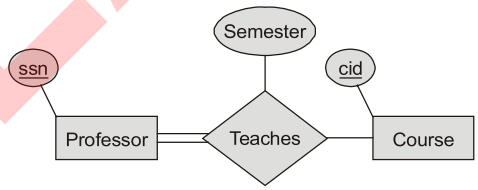
(b) Describe all data that is actually stored in the database.

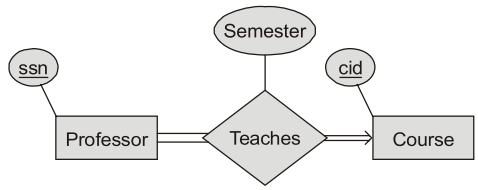
(c) Summarize flow the relation described in the conceptual schema are actually on disk.

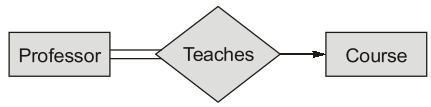
(d) All the above

Q.31 Which of the following ER diagram represents given requirement?
"Every professor teaches exactly one course and Every course must be taught by some professor".

(a)


(b)


(c)


(d)


Q.32 Which of the following statement true for physical schema of the levels of abstraction?

(a) Allows data access to be customized at the level of individual users.

(b) Describe all data that is actually stored in the database.

(c) Summarize how the relation described in the conceptual schema are actually on disk.

(d) All the above

Q.33 Which of the following is not the responsible of DBA?

(a) Design the logical and physical schemas, as well as widely used pointers of the external schema.

(b) Security and authorization.

(c) Concurrency control of the two or more concurrent execution of transaction.

(d) Data availability and recovery from failure.

Q.34 Consider the two sets of FD's for the relation R(A, B, C, D, E)

$F_1 = \{A \rightarrow B, AB \rightarrow C, D \rightarrow AC, D \rightarrow E\}$

$F_2 = \{A \rightarrow BC, D \rightarrow AE\}$

Which of the following statement true about FD set?

(a) $F_1 \subset F_2$ (b) $F_1 \supset F_2$

(c) $F_1 \equiv F_2$ (d) None of these

Q.35 Relation $R(A_1, A_2, A_3 \ldots A_n)$ n attributes on this relation m simple candidate key ($m \leq n$). How many super keys possible in this relation?

(a) $2^m - 1$ (b) $(2^n - 1) \cdot 2^{n-m}$

(c) $2^n - 1$ (d) $(2^m - 1)2^{n-m}$

Numerical Answer Type Questions

Q.36 Relation R(A1, A2, A3, A4,..., An). How many super keys in R with candidate keys

(a) {A1} (b) {A1A2, A3A4}

(c) {A1A2, A1A3} (d) {A1, A2, A3}

(e) {A1A2, A2A3, A3A4}

(f) {A1A2, A2A3A4, A5}

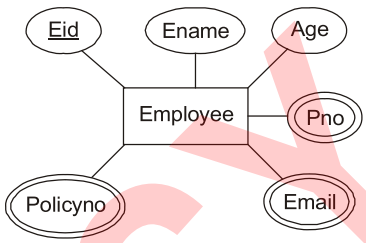
(g) {Every attribute of relation R}

Q.37 Relation R(A B C D E F). How many super keys in R if {A, BC, CD} candidate keys.

Q.38 Let E_1 and E_2 be two in an E/R diagram with simple single-valued attributes R_1 and R_2 are two relationship between E_1 and E_2, where R_1 is one-to many and R_2 is many-to-many. R_1 and R_2 do not have any attributes of their own. What is the minimum number of tables required to represent this situation in the relational model?

[GATE-2005]

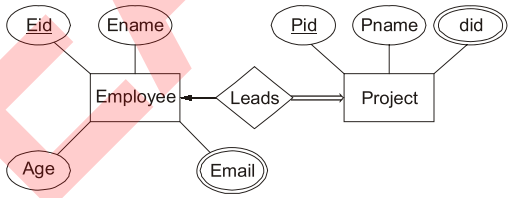
Q.39 Consider the following entity set



(*i*) How many minimum tables required which satisfy 1 NF?

(*ii*) How many minimum tables required which satisfy BCNF?

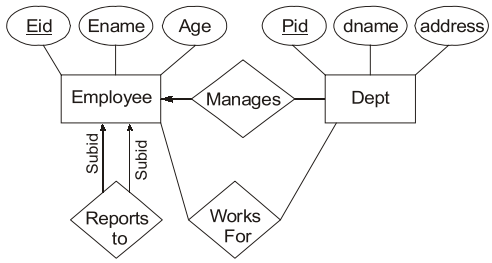
Q.40 Consider the following ER diagram:



(*i*) How many minimum RDBMS tables required for above ER diagram which satisfy 1 NF

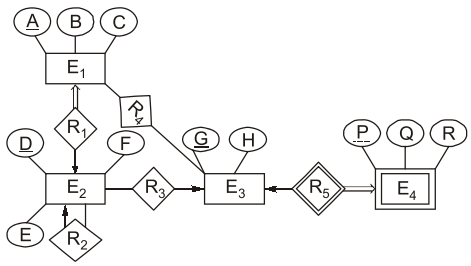
(*ii*) How many minimum RDBMS tables required which satisfy BCNF?

Q.41



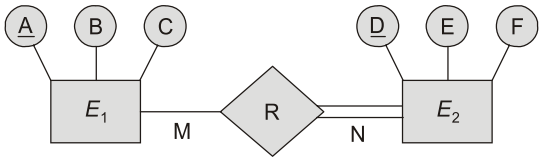
How many RDBMS tables required which satisfy 3NF?

Q.42 Consider the following ER diagram:



(*i*) How many minimum RDBMS tables required for above ER diagram
(*ii*) How many foreign keys required for minimized RDBMS table of above ER diagram
(*iii*) Total number of attribution in the minimized RDBMS tables of above ER diagram?

Q.43 Consider the following ERD:

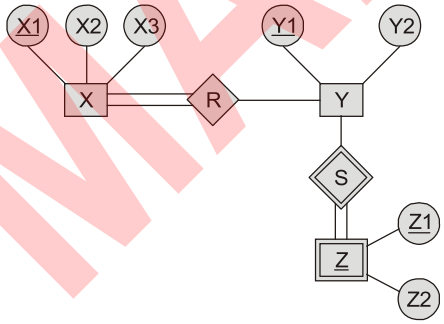


How many minimum relations which satisfy 1NF? [partial participation between E_1 and R should not lost in RDBMS design]

Multiple Select Questions

Q.44 Consider a relation scheme R = (A, B, C, D, E, F). The functional dependencies on R is {A → B, BC → D, D → A, E → C}. Which of the following is true?
(a) AEF is candidate key of R.
(b) BEF is candidate key of R.
(c) DEF is candidate key of R.
(d) BCF is candidate key of R.

Q.45 Consider the following ER diagram:



Which of the following is correct when ER-diagram is converted into minimum tables?
(a) Minimum 4 tables are required to represent X, Y, Z, R and S
(b) The one of the table will be {X1, X2, X3, Y1} where X1 is primary key and Y1 references Y
(c) The one of the table will be {Y1, Z1, Z2} where Y1 is primary key
(d) {Y1, Y2} is one of the table where Y1 is primary key.

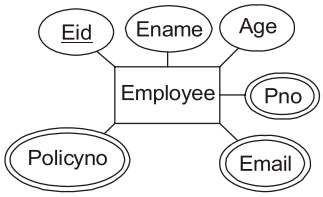
Q.46 Given the basic ER and relational modes which of the following are correct?
(a) An attribute of an entity can have more than one value.
(b) An attribute of an entity can be composite.
(c) In a row of relational table an attribute can have more than one value.
(d) In a row of a relational table an attribute can have exactly one value or a null value.

Q.47 Which of the following are correct?
(a) A relation R with schema (X, Y) satisfies the functional dependency X → Y . The tuples (2, 3) and (3, 3) both can be in R simultaneously
(b) Domain calculus is based on non procedural language.
(c) Relation R with an associated set of functional dependencies F is decomposed into BCNF. The redundancy due to functional dependencies in the resulting set of relation is zero
(d) A relation R(A, B, C, D, E) with FD's A → D, B → C, D → E, CE → B. If we project R on schema (ABC) then {AB} is the only candidate key.

Q.48 Which of the following is/are correct?
(a) An Entity set is a set of entities of the same type that share the same properties, or attributes.
(b) The attribute AGE is calculated from DATE_OF_BIRTH. The attribute AGE is Composite.
(c) Not applicable condition can be represented in relation entry as NULL.
(d) Data integrity constraints are used to Improve the quality of data entered for a specific property.

Try Yourself

T1. Consider the following entity set



(*i*) How many minimum tables required which satisfy 1 NF?

(a) 1 (b) 2
(c) 3 (d) 4

[Ans: (a)]

(*ii*) How many minimum tables required which satisfy BCNF?

(a) 1 (b) 2
(c) 3 (d) 4

[Ans: (d)]

T2. Consider a relation R(A, B, C, D, E) with the following functional dependencies:

$$ABC \rightarrow DE \text{ and } D \rightarrow AB$$

The number of superkeys of R is:

(a) 2 (b) 7
(c) 10 (d) 12

[Ans: (c)]

T3. DML is provided for

(a) Description of logical structure of database
(b) Addition of new structures in the database system
(c) Manipulation and processing of database
(d) Definition of physical structure of database system

[Ans: (c)]

T4. Which of the following is not part of Data Definition Language (DDL)

1. Deleting relations.
2. Defining.
3. Specifying the security and authorization for relation.
4. Inserting tuples, deleting tuples.
5. Physical storage structure information.

(a) 3 and 5
(b) Both 1 and 4
(c) 4 only
(d) 3 and 2

[Ans: (c)]

T5. View in a database system are important because:

1. They improve the efficiency of query execution.
2. They help provide data independence.
3. They allow the schema to change without forcing existing applications to be recompiled.
4. They help with access control by allowing users to see only a particular subset of the data in the database.

(a) 2 and 4 (b) 1 and 4
(c) 1 and 3 (d) 2 and 3

[Ans: (a)]

T6. The following two questions refer to the relational schema R(A, B, C, D, E, F, G, H) and the following functional dependencies over R:

$$A \rightarrow BCD,\ AD \rightarrow E,\ EFG \rightarrow H,\ F \rightarrow GH$$

Q.1 Based on the functional dependencies, there is one key for R. What is it?

[Ans: (AF)]

Q.2 One of the four functional dependencies can be removed without altering the key. Which one?

[Ans: ($EFG \rightarrow H$)]

T7. A database of research articles in a journal uses the following schema.

(VOLUME, NUMBER, STARTPAGE, ENDPAGE, TITLE, YEAR, PRICE)

The primary key is (VOLUME, NUMBER, STARTPAGE, ENDPAGE) and the following functional dependencies exist in the schema.

(VOLUME, NUMBER, STARTPAGE, ENDPAGE) → TITLE

(VOLUME, NUMBER) → YEAR

(VOLUME, NUMBER, STARTPAGE, ENDPAGE) → PRICE

The database is redesigned to use the following schemas.

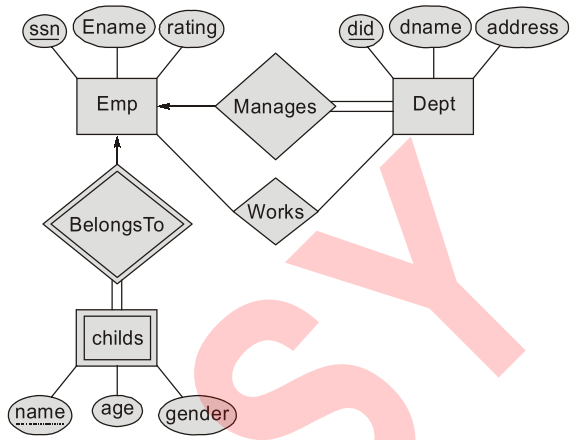
(VOLUME, NUMBER, STARTPAGE, ENDPAGE, TITLE, PRICE)
(VOLUME, NUMBER, YEAR)

Which is the weakest normal form that the new database satisfies, but the old one does not?

(a) 1NF (b) 2NF
(c) 3NF (d) BCNF

[GATE-2016, Ans: (b)]

T8. Consider the given ERD:



How many relational tables for given ERD?

[Ans: (4)]

T9. Consider the following relational schema:
R(A, B, C, D, E) how many superkeys in relation R if every two attributes of relation R is candidate keys?

[Ans: (26)]

■■■■

2 Normalization

Multiple Choice Questions

Q.1 The following functional dependencies are given: AB → CD, AF → D, DE → F, C → G, F → E, G → A. Which one of the following options is false?

(a) $\{CF\}^+ = \{ACDEFG\}$
(b) $\{BG\}^+ = \{ABCDG\}$
(c) $\{AF\}^+ = \{ACDEFG\}$
(d) $\{AB\}^+ = \{ACDFG\}$

[GATE-2006]

Q.2 Consider a relation scheme R = (A,B,C,D,E,H) on which the following functional dependencies hold: {A → B, BC → D, E → C, D →A}. What are the candidate keys of R?

(a) AE, BE
(b) AE, BE, DE
(c) AEH, BEH, BCH
(d) AEH, BEH, DEH

[GATE-2005]

Q.3 Which of the following statement false if relation R is in 3NF but not BCNF?

(a) R must consist at least two over lapped candidate keys.
(b) R must consist proper subset of candidate key determines proper subset of some other candidate key.
(c) R must consist at most one compound candidate key other candidate keys simple candidate key.
(d) R must consist at least two compound candidate keys.

Q.4 Consider the following declaration:
F : set of functional dependencies
R : relation of functional dependencies
R : (A, B, C, D, E)
F : {A → BC;
CD → E;
B → D;
E → A}
Then candidate keys for R is _____.

(a) A, E, CD, BC (b) AB, EB, CD, BC
(c) A, E, C, BD (d) A, E, CD, BD

Q.5 Consider the following relation R(TUVWXYZ) and FD set {TU → VW, TY → W, WX → Y, V → Z, Y → X, Z → T}
How many number of candidate keys in relation R?

(a) 3 (b) 4
(c) 5 (d) 6

Q.6 Let R be an RDBMS with attributes A_1, A_2, ..., A. Let S denotes the set $\{A_1, A_2, ..., A_n\}$. Let T ⊆ S be a set of attributes that forms a candidate key. Then which of the following is/ are True?

P : T → S – T
Q : ∃P ⊂ T s.t. P → S – P
R : ∀ Q ⊇ T s.t. Q → S – Q

(a) Only P is true (b) P and Q are true
(c) P and R are true (d) Q and R are true

Q.7 Which of the following functional dependencies implied in the given FD set {AB → CD, AF → D, DE → F, C → G, F → E, G → A}.

(a) AB → F (b) BG → E
(c) CF → D (d) AF → B

Q.8 If the set of functional dependencies F = {A → BC, CD → E, E → C, D → AEH, ABH → BD, DH → BC}, then what is the canonical cover of F?
(a) {A → BC, C → E, E → C, D → AE, DH → B}
(b) {A → BC, E → C, D → AEH, AH → D}
(c) {A → B, E → C, D → AH, AH → D}
(d) {A → C, C → E, E → C, D → AH, D → BC}
[JNUEE-2007]

Q.9 Consider two sets of functional dependencies F and G if F = {A → C, AC → D, E → AD, E → H} then the equivalent G is
(a) G = {A → CD, E → AH}
(b) G = {A → CH, E → ADH}
(c) G = {A → CD, E → H}
(d) G = {A → AH, E → CH}

Q.10 Identify minimal cover for the following FD set {AB → C, C → A, BC → D, ACD → B, BE → C, EC → FA, CF → BD, D → E}
(a) {AB → C,C → ABD, BE → C,EC → AF, D→E}
(b) {A → C, C → ABD, BE → C, EC → F, D→E}
(c) {A → C, C → ABD, BE → C, EC→AF, D→E}
(d) {AB → C, C → A, BC → D, BE → C, EC→F, D → E, CF → B}

Q.11 Consider the following relational schema:
R(ABCDEGH) with FD set {AB → C, AC → B, AD → E, B → D, BC → A, E → G}
The following relations are sub relations of the relation R.
(i) R_1(ABC)
(ii) R_2(ABCD)
(iii) R_3(ABCEG)
How many sub relations of R in BCNF?
(a) 1 (b) 2
(c) 3 (d) 0

Q.12 Test the following decompositions of relations lossless join decomposition and dependency preserving decomposition satisfied or not?
(a) R(ABCDEFGHIJ) and FD sets {AB → C, A → DE, B → F, F → GH, D → IJ}
(*i*) D_1 = {DIJ, ADE, FGH, BF, ABC}
(*ii*) D_2 = {DIJ, ACE, FGH, BF, ADC}
(*iii*) D_3 = {FGH, DIJ, ADEBF, ABC}
(b) R(ABCDEGH) and FD's sets {AB → C, AC → B, AD → E, B → D, BC → A, E → G}
(*i*) D_1 = {AB, BC, ABDE, EG}
(*ii*) D_2 = {ABC, ACDE, ADG}
(c) R(ABCDEG) and FD's sets {AB → C, AC → B, AD → E, B → D, BC → A, E → G}
(*i*) D_1 = {AB, BC, ABDE, EG}
(*ii*) D_2 = {ABC, ACDE, ADG}
(d) R(ABCDEF) and FD's sets F = {A → B, AC → DE, BD → F}
Decomposed into {AB, BDF, ACDE}

Q.13 Let R (A, B, C, D) be a relational schema wit the following functional dependencies: A → B, B → C, C → D and D → B. The decomposition of R into (A, B), (B, C) and (B, D)
(a) gives a lossless join, and is dependency preserving
(b) gives a lossless join, but is not dependency preserving
(c) does not give a lossless join, but is dependency preserving
(d) does not give a lossless join and is not dependency preserving
[GATE IT-2008]

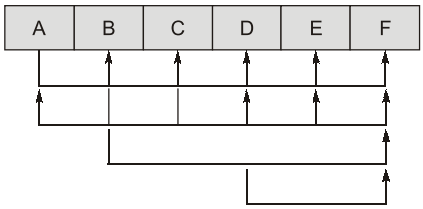
Q.14 Let R be a relationship schema and let F be a set of functional dependencies on R. Let R_1 and R_2 form a decomposition of R, F^+ must contain
(a) $R_1 \cup R_2 \to R_1$ or $R_1 \cup R_2 \to R_2$
(b) $R_1 \cup R_2 \to R_1$ or $R_1 \cap R_2 \to R$
(c) $R_1 \cup R_2 \to R_2$ or $R_1 \cap R_2 \to R$
(d) $R_1 \cap R_2 \to R_1$ or $R_1 \cap R_2 \to R_2$
[DRDO-2009]

Q.15 For a relational schema R, K_1 and K_2 are the only candidate keys. R has a functional dependency X → A where X ia set of attributes and A is an attribute. It is known that $A \in K_1$ and $A \notin K_2$ and X is not a superkey. Which of the following is true?
(a) R could be in BCNF
(b) R is surely not in BCNF, but could be in 3NF
(c) R is surely not in 3NF, but could be in 2NF
(d) R is surely not in 2NF, but could be in 1NF
[ISRO-2009]

Q.16 A relation with 2 attributes is always in
(a) 4NF (b) BCNF
(c) 5NF (d) DK/NF
[DRDO-2009]

Q.17 An RDBMS with 3 attributes has the following functional dependencies A → B A → C C → B
(a) 2NF (b) 3NF
(c) 4NF (d) BCNF
[DRDO-2009]

Q.18 The best normal form the instance is



(a) 1NF (b) 2NF
(c) 3NF (d) BCNF

Q.19 Consider the relation r (A, B, C, D, E) and the set F = {AB → CE, E → AB, C → D}. What is the highest normal form of this relation?
(a) 1NF (b) 2NF
(c) 3NF (d) BCNF

Q.20 Consider the following relational schemas for a library database.
Book (Title, Author, Catalog_no, Publisher, Year, price)
Collection (Title, Author, Catalog_no)
with the following functional dependencies:
1. Title, Author → Catalog
2. Catalog_no → Title, Author, Publisher, Year
3. Publisher, Title, Year → Price

Assume {Author, Title} is the key for both schemes: which of the following statements is true?
(a) Both Book and Collection are in BCNF
(b) Both Book and Collection are 3NF only
(c) Book is in 2NF and Collection is in 3NF
(d) Both Book and Collection are in 2NF only
[GATE-2008]

Q.21 A relation Empdtl is defined with attribute empcode (unique), name, street, city, state and pincode. For any pincode, there is only one city and state. Also, for any given street, city and state, there is just one pincode. In normalization terms, Empdtl is a relation in
(a) 1NF only
(b) 2NF and hence also in 1NF
(c) 3NF and hence also in 2NF and 1NF
(d) BCNF and hence also in 3NF, 2NF an 1NF
[GATE IT-2004]

Q.22 A table has fields F1, F2, F3, F4, F5 with the following functional dependencies

F1 → F3, F2 → F4, (F1. F2) → F5

In terms of Normalization, this table is in
(a) 1NF (b) 2NF
(c) 3NF (d) None of these
[GATE IT-2005]

Q.23 Let R (A, B, C, D, E, P, G) be a relational schema in which the following functional dependencies are known to hold: AB → CD, DE → P, C → E, P → C and B → G. The relational schema R is
(a) in BCNF
(b) in 3NF, but not in BCNF
(c) in 2NF, but ont in 3NF
(d) not in 2NF
[GATE IT-2008]

Q.24 The relation scheme student performance (name, courseno, rollno, grade) has the following functional dependencies
name, courseno → grade
rno, courseno → grade
name → rollno
rollno → name
The highest normal form of this relation scheme is
(a) 2NF (b) 3NF
(c) BCNF (d) 4NF

Q.25 Consider a relation with schema R(A, B, C, D) and functional dependencies {AB → C, C → D, D → A} what is the highest normal form?
(a) 1NF (b) 2NF
(c) 3NF (d) BCNF

Q.26 What is the highest normal form of the relation R(A, B, C, D, E) with functional dependencies {AB → DE, C → E, D → C, E → A}
(a) 1NF (b) 2NF
(c) 3NF (d) BCNF

Q.27 Which of the following relation is in BCNF?
(a) R(A, B, C, D) with FD's {AB → C, BC → D, CD → A, AD → B}
(b) R(A, B, C, D, E) with FD's {AB → C, DE → C, B → D}
(c) R(A, B, C, D, E) with FD's {AB → C, C → D, D → B}
(d) None of these

Q.28 Consider the following relational schema R(ABCDEFG) with FD set {AB → C, BC → A, B → D, D → E}
How many minimum relations required to decompose R into BCNF which satisfy lossless join and dependency preserving decompositions?
(a) 3 (b) 2
(c) 4 (d) 5

Q.29 Consider the following functional dependencies in a database:
Data_of_Birth → Age
Age → Eligibility
Name → Roll_number
Roll_number → Name
Course_number → Course_name
Course_number → Instructor
(Roll_number; Course_number) → Grade
The relation (Roll_number; Name, Date_of_birth, Age) is
(a) in 2NF but not in 3NF
(b) in third normal form but not in BCNF
(c) in BCNF
(d) in none of the above

[GATE-2003]

Q.30 The relation scheme student Performance (name, courseNo, rollNo, grade) has the following functional dependencies:
name, courseNo → grade
RollNo, courseNo → grade
name → rollNo
rollNo → name
The highest normal form of this relation scheme is
(a) 2NF (b) 3NF
(c) BCNF (d) 4NF

[GATE-2004]

Linked Answer for Q.31 & Q.32
Relation R has eight attributes ABCDEFGH. Fields of R contain only atomic values.
F = {CH → G, A → BC, B → CFH, E → A, F → EG} is a set of functional dependencies (FDs) so that F is exactly the set of FDs that hold for R.

Q.31 How many candidate keys does the relation R have?
(a) 3 (b) 4
(c) 5 (d) 6

[GATE-2013]

Q.32 The relation R is
(a) in 1NF, but not in 2NF
(b) in 2NF, but not in 3NF
(c) in 3NF, but not in BCNF
(d) in BCNF

[GATE-2013]

Q.33 Consider the relation R = {A, B, C, D, E, F, G, H, I, J} and set of FDS are F = {AB → C, A → DE, B → F, F → GH, D → IJ} if we decompose R into 3NF, then possible tables are
(a) {D, I, J} {A, D, E} {F, G, H} {B, F} {A, B, C}
(b) {D, I, J} {A, C, E} {F, G, H} {B, F} {A, D, C}
(c) {F, G, H} {D, I, J} {A, D, E, B, F} {A, B, C}
(d) None of these

Q.34 R(A, B, C, D) is a relation, which of the following doesn't have a loss less join dependency preserving BCNF decomposition?
(a) A → B, B → CD
(b) A → B, B → C, C → D
(c) AB → C, C → AD
(d) A → BCD

Q.35 Consider the schema R = (S T U V) and the dependencies S → T, T → U, U → V and V → S. Let $R = (R_1$ and $R_2)$ be a decomposition such that $R_1 \cap R_2 = \phi$ then the decomposition is

(a) not in 2 NF
(b) in 2NF but not 3 NF
(c) in 3 NF but not 2 NF
(d) in both 2 NF and 3 NF

Q.36 Relation R with an associated set of functional dependencies, F is decomposed into BCNF. The redundancy (arising out of functional dependencies) in the resulting set relations is
(a) Zero
(b) More than zero but less than that of an equivalent 3NF decomposition
(c) Proportional to the size of F^+
(d) Indeterminate

[GATE-2002]

Q.37 Relation R is decomposed using a set of functional dependencies, F and relation S is decomposed using another set of functional dependencies G. One decomposition is definitely BCNF, the other is definitely 3NF, but it is not know which is which. To make a guaranteed identification, which one of the following tests should be used on the decomposition? (Assume that the closures of F and G are available).
(a) Dependency-preservation
(b) Lossless-join
(c) BCNF definition
(d) 3NF definition

[GATE-2002]

Q.38 Which-one of the following statements about normal forms is FALSE?
(a) BCNF is stricter than 3NF
(b) Loss less, dependency-preserving decomposition into 3NF is always possible
(c) Loss less, dependency-preserving decomposition into BCNF is always possible
(d) Any relation with two attributes is BCNF

[GATE-2005]

Q.39 A given relation is known to be in third normal form. Select the statement which can be inferred from this.
(a) All attributes contribute to the primary key
(b) Each non-key attribute determine the primary key
(c) Each non-key attribute is determine by the primary key
(d) Every determinant is a candidate key

Q.40 Consider the following relation R(ABCDE) with functional dependency set {A → BC, CD → E, B → D, E → A} decomposed relation R into R_1(ABC) R_2(CDE) which of the statement true about decomposition?
(a) Lossless join decomposition and dependency preserving decomposition.
(b) Lossless join decomposition but not dependency preserving decomposition.
(c) Lossy join decomposition and dependency preserving decomposition.
(d) Lossy join decomposition and not dependency preserving decomposition.

Q.41 Consider the following relation R(ABC) with FD set {A → B, C → B, C → B} relation R decomposed into
(i) R_1(AB) R_2(BC)
(ii) R_1(AC) R_2(BC)
(iii) R_1(AB) R_2(AC)
How many of the above decompositions are lossless join decompositions?
(a) 1 (b) 2
(c) 2 (d) 0

Q.42 Consider the following relation R(ABC) with FD set {A → B, C → B, C → B} relation R decomposed into
(i) R_1(AB) R_2(BC)
(ii) R_1(AC) R_2(BC)
(iii) R_1(AB) R_2(AC)
How many of the above decompositions are dependency preserving decompositions?
(a) 1 (b) 2
(c) 3 (d) 0

Q.43 Which of the following relation can decompose into BCNF with dependency preserving and lossless join decomposition.

(i) R(ABCDE)
{AB → C, C → AB, C → D, D → E}

(ii) R(ABCDE)
{AB → C, C → A, C → D, D → E}

(a) Only (i)
(b) Only (ii)
(c) Both (i) and (ii)
(d) None of these

Q.44 Consider the following relational schema R(ABCD). Which of FD set of R does not have BCNF decomposition [i.e. not possible to decompose BCNF with lossless join and dependency preserving decomposition].

(i) {B → C, D → A}
(ii) {ABC → D, D → A}
(iii) {AB → CD, C → A, D → B}
(a) (ii) only
(b) (ii) and (iii)
(c) (i) and (ii)
(d) (i), (ii) and (iii)

Common Data Questions for Q.45 to Q.46:

Consider the following relational schema R(A, B, C, D, E, F) and functional dependency set {AB → C, C → A, BC → D, ACD → B, BE → C, EC → AF, CF → BD, D → E}.

Q.45 How many candidate keys for relation R?
(a) 7
(b) 4
(c) 5
(d) 6

Q.46 What is highest NF satisfied by relation R?
(a) 1NF
(b) 2NF
(c) 3NF
(d) BCNF

Numerical Answer Type Questions

Q.47 Find minimal cover of the following functional dependencies
I. {A → BCDEF, BC → ADEF, B → F, D → E}
II. {AB → C, D → E, AB → E, E → C}
III. {AB → C, BC → A, A → BC, B → AC, C → AB}

Q.48 How many number of candidate keys in the following relation R(ABCDEH) with FD's {A→ BC, CD → E, E → C, D → AEH, ABH → BD, DH → BC}?

Q.49 Consider the following relation R(ABCDE) with FD's {A → C, BC → E, ED → A}. How many candidates keys in R?

Q.50 Consider the following relation R(ABCDEFG) with functional dependencies {AB → CD, AF → D, DE → F, C → G, F → E, G → A}. How many candidate keys in the above relation R?

Q.51 Consider the following relation R(ABCDEH) with functional dependencies {A → BC, CD → E, E → C, D → AEH, ABH → BD, DH → BC}. How many candidate keys of the above relation R?

Q.52 Relation R(A B C D E F G H I J) with FD set {AB → C, A → DE, B → F, F → GH, D → IJ}. How many minimum number of relations required for 2 NF lossless and dependency preserve decomposition?

Q.53 How many minimum relations required to decompose following relation R(ABCD) WITH FD's {A → B, C → D} with lossless join and dependency preserving BCNF decomposition?

Q.54 How many foreign keys required for the following relation R(ABCDEF) with FD's {AB→C, BC → A, AC → B, D → E} into BCNF and lossless join, dependency preserving decomposition?

Multiple Select Questions

Q.55 Which of the following options are correct?

(a) A prime attribute can transitively dependent on a key in a BCNF relation.

(b) A relation in which every key has only one attribute is in 2NF.

(c) A prime attribute can be transitively dependent on a key in a 3NF relation.

(d) There is always a decomposition into Boyce-code normal form that is lossless and dependency preserving.

Q.56 Relation R = (L, M, N, O, P) where the dependencies M → O, NO → P, P → L and L → MN holds.

Assume R is decomposed into R_1(M, O) and R_2(L, M, N, P). Then which of the following is correct?

(a) The above decomposition is lossless join.

(b) The highest normal form satisfied by the above decomposition is BCNF.

(c) The above decomposition does not satisfied dependency preserving as O → P does not preserved.

(d) The above decomposition does not satisfy dependency preserving as ON → P does not preserved.

Q.57 Relation R is decomposed using a set of functional dependencies F and relation S is decomposed using another set of functional dependencies G. One of the decomposition is 3NF and another is BCNF. Assume the closure of F and G are available but if not clear which relation is in which normal forms. Which of the following is true?

(a) Using BCNF definition we can certainly tell about the normal forms of F and G.

(b) Using dependency preservation we may/ may not tell about the normal forms of F and G.

(c) Using lossless join we can't tell about the normal forms of F and G.

(d) Using 2NF definition we can find out the normal forms of F and G

Q.58 Which-one of the following statements about normal forms is TRUE?

(a) BCNF is stricter than 3NF.

(b) Lossless decomposition and dependency preserving in 3NF is always possible.

(c) Lossless decomposition dependency preserving in BCNF is always possible.

(d) Any relation with two attributes is BCNF.

Q.59 Empdt1(empcode, name, street, city, state, pincode).

For any pincode, there is only one city and state. Also, for given street, city and state, there is just one pincode. In normalization terms, empdt1 is a relation in

(a) 1NF (b) 2NF

(c) 3NF (d) BCNF

Try Yourself

T1. Find highest normal form of the following relations:

(a) R(ABCD) {AB → C, C → D, D → A}

(b) R(ABCD) {B → C, B → D}

(c) R(ABCD) {AB → C, BC → D, CD → A, AD → B}

(d) R(ABCD) {A → B, B → C, C → D, D → A}

(e) R(ABCDE) {AB → C, DE → C, B → D}

(f) R(ABCDE) {AB → C, C → D, B → D, D → E}

(g) R(ABCDE) {A → B, B → C, C → D, D → A}

T2. Decompose following relations into 2NF, 3NF BCNF and state what are relations failed to decompose into BCNF and dependency preserving decomposition.

(a) R(ABCD) {AB → C, C → D, D → A}

(b) R(ABCD) {B → C, B → D}

(c) R(ABCD) {AB → C, BC → D, CD → A, AD → B}
(d) R(ABCD) {A → B, B → C, C → D, D → A}
(e) R(ABCDE) {AB → C, DE → C, B → D}
(f) R(ABCDE) {AB → C, C → D, D → B, D → E}

T3. The purpose of schema normalization is to:
1. Eliminate redundant data stored in the database.
2. Reduce the number of joins required to satisfy a query.
3. Reduce the number of anomalies that can occur during inserts, deletes, and updates.
4. Convert the data to a canonical form to promote schema integration.

(a) 1 and 4
(b) 2 and 3
(c) 1, 2 and 3
(d) 1 and 3

[Ans : (d)]

T4. Consider the following FD set
{AB → C, C → A, BC → D, ACD → B, BE → C, EC → AF, CF → BD, D → E}
Which of the following FD set is minimal cover of given FD set?

(a) {AB → C, C → A, BC → D, BE → C, EC → F, CF → D, D → E}
(b) {AB → C, C → A, BC → D, BE → D, EC → F, CF → B, D → E}
(c) {AB → C, C → A, CD → B, EC → F, CF → B, D → E}
(d) {AB → C, C → A, BC → D, BE → C, EC → A, CF → B, D → E}

[Ans : (b)]

T5. Consider the following relational schema:
R(A, B, C, G, H, I) with FD's {A → B, A → C, CG → A, CG → I, B → H}
How many candidate keys of relation R?

[Ans : (2)]

■■■■

3 Relational Algebra, Tuple Relational Calculus and SQL

Multiple Choice Questions

Q.1 Consider the following DB table Emp (Eid, Ename, gender, salary, dno)
(a) Retrieve Eid's who gets highest salary
(b) Retrieve Eid's who gets 2nd highest salary
(c) Retrieve Eid's who gets highest salary for each department
(d) Retrieve dept number for which at least two employees

Q.2 Which of the following relational algebraic operation is not a commutative operation?
(a) Union (b) Intersection
(c) Selection (d) Projection

Q.3 With the help of which of the following relations operation set we can perform division on relations?
(a) $\{\pi, \times, -\}$ (b) $\{\sigma, \times, -\}$
(c) $\{\times, -\}$ (d) $\{\times\}$

Q.4 Let r and s be two relations over the relation schemes R and S respectively, and let A be an attribute in R. Then the relational algebra expression $\sigma_{A=a}(r \bowtie s)$ is always equal to
(a) $\sigma_{A=a}(r)$ (b) r
(c) $\sigma_{A=a}(r) \bowtie s$ (d) None of these

[GATE-2001]

Q.5 Information about a collection of students is given by the relation studInfo (studId, name, sex). The relation enroll (studId, CourseId) gives which student has enrolled for (or taken) what course(s). Assume that every course is taken by at least one male and at least one female student. What does the following relational algebra expression represent?

$\pi_{courseId}((\pi_{studId}(\sigma_{sex = \text{"female"}}(studInfo))) \times \pi_{courseId}(enroll) - enroll)$

(a) Courses in which all the female students are enrolled
(b) Courses in which a proper subset of female students are enrolled
(c) Courses in which only male students are enrolled
(d) None of the above

[GATE-2007]

Q.6 Let R and S be two relations with the following schema
R(P, Q, R1, R2, R3)
S(P, Q, S1, S2)
where {P, Q} is the key for both schemas. Which of the following queries are equivalent?

1. $\Pi_P(R \bowtie S)$
2. $\Pi_P(R) \bowtie \Pi_P(S)$
3. $\Pi_P(\Pi_{P,Q}(R) \cap \Pi_{P,Q}(S))$
4. $\Pi_P(\Pi_{P,Q}(R) - (\Pi_{P,Q}(R) - \Pi_{P,Q}(S)))$

(a) Only 1 and 2 (b) Only 1 and 3
(c) Only 1, 2 and 3 (d) Only 1, 3 and 4

[GATE-2008]

Q.7 Let R_1(A, B, C) and R_2(D, E) be two relation schema, where the primary keys are shown underlined, and let C be a foreign key in R_1 referring to R_2. Suppose there is no violation of the above referential integrity constraint in the corresponding relation instances r_1 and r_2. Which one of the following relational algebra expressions would necessarily produce an empty relation?

(a) $\Pi_D(r_2) - \Pi_C(r_1)$
(b) $\Pi_C(r_1) - \Pi_D(r_2)$
(c) $\Pi_D(r_1 \bowtie_{C \neq D} R_2) - \Pi_C(r_1)$
(d) $\Pi_C(r_1 \bowtie_{C = D} R_2)$

[GATE-2004]

Q.8 Consider the following relational schema pertaining to a student database

Student (rollno, name, address)
Enroll (rollno, courseno, coursename)

Where the primary keys are shown underlined. The number of tupelos in the student and Enroll tables are 120 and 8 respectively. What are the maximum and minimum number of tuples that can be present in (Student*Enroll), where '*' denotes natural join?

(a) 8, 8 (b) 120, 8
(c) 960, 8 (d) 960, 120

Q.9 Suppose the adjacency relation of vertices in a graph is represented in a table Adj (x, y). Which of the following queries cannot be expressed by a relational algebra expression of constant length?

(a) List of all vertices adjacent to a given vertex
(b) List all vertices which have self loops
(c) List all vertices which belong to cycle of less than three vertices
(d) List all vertices reachable from a given vertex

Q.10 Given the relations

Employee (name, salary, dept no) and
Dept (dept no, dept name, address)

Which of the following queries cannot be expressed using the basic relational algebra operations (σ, π, x, $\bowtie$, $\cup$, $\cap$, —)?

(a) Department address of every employee
(b) Employees whose name is the same as their department name
(c) The sum of all employees salaries
(d) All employees of a given department

Q.11 Consider the relation Student(name, sex, marks) where the primary key is shown underlined, pertaining to student in a class that has at least one boy and one girl. What does the following relational algebra expression produce

Note: ρ is the rename operator)

$\pi_{name}(\sigma_{sex=female}(Student) - \pi_{name}(Student \bowtie_{(sex = female \wedge x = male \wedge marks \leq m)} \sigma_{n,x,m}(Student)))$

(a) names of girl student with the highest marks
(b) names of girl student with more marks than some boy student
(c) names of girl student with marks not less than some boy student
(d) names of girl students with more marks than all the boy students

Q.12 Consider the following relational schema Adjacency(x, y) used to store edges of the graph. Which of the following statement not possible to represent constant length of relational algebra?

(a) Retrieve vertices with self loop in graph.
(b) Retrieve vertices which forms loop in graph.
(c) Retrieve vertices whose indegree "0".
(d) Retrieve vertices which are reachable two edge path length from given vertex.

Q.13 Consider the relational schema given below, where **eId** of the **dependent** is a foreign key referring to **empId** of the relation **employee**. Assume that every employee has at least one associated dependent in the **dependent** relation.

employee (empId, empName, empAge)
dependent(depId, eId, depName, depAge)

Consider the following relational algebra query:

$\Pi_{empId}(employee) - \Pi_{empId}(employee \bowtie_{(empId=eId) \wedge (empAge \leq depAge)} dependent)$

The above query evaluates to the set of empIds of employees whose age is greater than that of

(a) some dependent.
(b) all dependents.
(c) some of his/her dependents.
(d) all of his/her dependents.

[GATE-2014 (Set-3)]

Common Data For Q.14 & Q.15

Consider the following relational database schema:

employee (emp no, name, address)
project (p no. p_name)
work_on(emp no, p no)

Q.14 We have a relational algebra expression on the above schema:

$\Pi_{name}(employee) - \Pi_{name}(employee \otimes work\_on)$

Here $\otimes$ denotes natural join operator. Then which of the following query best resembles the above relationship algebra expression?

(a) Find the name of all employees working in a project.
(b) Find the name of all employees working in all projects.
(c) Find the name of all employees who don't work in all projects.
(d) Find the name of all employees who don't work in any project.

[DRDO-2009]

Q.15 Find all addresses of employees working in the project with p_name = "database". Which of the following SQL represents the above query.

(a) SELECT address FROM employee, project WHERE (p_name= "database") ^ (work_on.emp_no=employee.emp_no)
(b) SELECT address FROM employee, project, work_on WHERE (p_name="database") ^ (work_on.emp_no=employee.emp_no) ^ (work_on.emp_no=employee.emp_no) ^ (work_on.p_no-project.p_no)
(c) SELECT address FROM employee, work_on WHERE (p_name="database") ^ (work_on.emp_no=employee.emp_no)
(d) SELECT address FROM project, work_on WHERE (p_name="database") ^ (work_on.emp_no=employee.emp_no)

[DRDO-2009]

Q.16 In SQL, relations can contain null values, and comparisons with null values are treated as unknown. Suppose all comparisons with a null value are treated as false. Which of the following pairs is not equivalent?

(a) $x = 5$ not (not ($x = 5$))
(b) $x = 5$ $x > 4$ and $x < 6$, where x is an integer
(c) $x \neq 5$ not ($x = 5$)
(d) None of the above

[GATE-2000]

Q.17 Consider the set of relations shown below and the SQL query that follow:

Students: (Roll_number, Name, Date_of_birth)
Courses: (Course_number, Cours_name, Instructor)
Grades: (Roll_number, Course_number, Grade)

```
select distinct Name
from Students, Courses, Grades
where Students. Roll_number = Grades. Roll_number
and Courses. Instructor = Korth
and Courses. Course_number = Grades. Course_number
and Grades. grade = A
```

Which of the following sets is computed by the above query?

(a) Names of students who have got an A grade in all courses taught by Korth
(b) Names of students who have got an A grade in all courses
(c) Names of students who have got an A grade in at least one of the courses taught by Korth
(d) None of the above

[GATE-2003]

Q.18 The relation book (title, price) contains the titles and prices of different books. Assuming that no two books have the same price, what does the following SQL?

```
Select title
from book as B
where (select count (*)
from book as T
where T. price> B. Price) < 5
```

(a) Titles of the four most expensive books
(b) Titles of the fifth most inexpensive book
(c) Titles of the fifth most expensive book
(d) Titles of the five most expensive books

[GATE-2005]

Q.19 Consider the relation account (customer, balance) where customer is a primary key and there are no null values. We would like to rank customers according to decreasing balance. The customer with the largest balance gets rank 1. Ties are not broken but ranks are skipped: if exactly two customers have the largest balance they each get rank 1 and rank 2 is not assigned.

Query 1: Select A. customer, count (B. customer)
from account A, account B
where A. balance < = B. balance
group by A. customer

Query 2: SelectA. customer, 1 + count (B. customer)
from account A, account B
where A.balance < B. balance
group by A. customer

Consider these statements about Query 1 and Query 2.

1. Query 1 will produce the same row set as Query 2 for some but not all database.
2. Both Query 1 and Query 2 are correct implementations of the specification.
3. Query 1 is a correct implementation of the specification but Query 2 is not.
4. Neither Query 1 nor 2 is a correct implementation of the specification.
5. Assigning rank with a pure relational Query takes less time than scanning in decreasing balance order and assigning ranks using ODBC.

Which two of the above statements are correct?

(a) 2 and 5 (b) 1 and 3
(c) 1 and 4 (d) 3 and 5

[GATE-2006]

Q.20 Consider the following database

Works (Ename, Comp-name, sex, salary) and the SQL query:

Select * W1. Ename

from Works W1, (Select Avg (sal) AS AVG-SAL, Comp-name from WORKS WHERE sex = 'Male' GROUP BY comp-name) W2

WhereW1 . comp-name = W2.comp-name AND W1. salary > W2.AVG-sal

What will be the output?

(a) Retrieves names of all employees who earn more than average salary of all Emp's of their company
(b) Retrieves names of all male employees who earn more than average salary of all employees of their company
(c) Retrieves names of all employees who earn more than average salary of all male employees of their company
(d) Retrieves names of all male employees who earn more than average salary of all male employees of their company

Q.21 Consider the relation Enrolled (SID, CID) in which (Sid, Cid) is the primary key, and the relation paid (Sid, amount) where sid is the primary key assume no null values and no foreign keys or integrity constraints.

Query 1: Select sid from Enrolled where sid in (select sid from paid)

Query 2: Select sid from paid where sid in (select sid from Enrolled)

Query 3: Select E.sid from Enrolled E, paid P where E.sid = p.sid

Query 4: Select sid from paid where exists (Select * from Enrolled where enrolled.sid = paid.sid)

Which one of the following statement is correct?

(a) All queries return identical row sets for any database
(b) Query 2 and Query 4 return identical row sets for all dabaase but there exist databases for which Query 1 and Query 2 return different row sets
(c) There exist databases for which Query 3 returns strictly fewer rows than Query 2
(d) There exist databases for which Query 4 will encounter an integrity violation at run time

Q.22 Consider the table Employee (EmpId, name, department, salary) and the two queries Q1, Q2 below:

Assuming that department 5 has more than one employee, and we want to find the employee who get higher salary than anyone in the department 5, which one of the statement is **true** for any arbitrary employee table?

Q1 Select e.empId
From employee e
where not exists(SELECT * from employee S where S.department = '5' and S.salary >= e.salary)

Q2 Select e.empID
from employee e
where e.salary > any
(SELECT distinct salary from employee S where S.department = 5)

(a) Q1 is the correct query
(b) Q2 is the correct query
(c) Both Q1 and Q2 produce the same answer
(d) Neither Q1 nor Q2 is the correct query

Q.23 A table T1 in a relational DB has the following rows and columns

Roll No.	Marks
1	10
2	20
3	30
4	NULL

The following sequence of SQL statements are executed on table T1

Update T1 set marks = marks + 5
Select Avg (marks) from T1

What is the output of the select statement

(a) 18.75 (b) 20
(c) 25 (d) NULL

Q.24 Given relations $r(w, x)$ and $s(y, z)$, the result of

select distinct w, x
from r, s

is guaranteed to be same as r, provided

(a) r has no duplicates and s is non-empty
(b) r and s have no duplicates
(c) s has no duplicates and r is non-empty
(d) r and s have the same number of tuples

[GATE-2000]

Q.25 Consider a database table T containing two columns X and Y each of type integer. After the creation of the table, one record (X = 1, Y = 1) is inserted in the table.

Let MX and MY denote the respective maximum values of X and Y among all records in the table at any point in time. Using MX and MY, new records are inserted in the table 128 times with X and Y values being MX + 1, 2*MY + 1 respectively. It may be noted that each time after the insertion, values of MX and MY change. What will be the output of the following SQL query after the steps mentioned above are carried out?

SELECT Y FROM T WHERE X = 7;

(a) 127 (b) 255
(c) 129 (d) 257

[GATE-2011]

Q.26 Consider the following relational schema:

employee(**empId**, empName, empDept)
customer(**custId**, custName, salesRepId, rating)

salesRepId is a foreign key referring to **empId** of the employee relation. Assume that each employee makes a sale to at least one customer. What does the following query return?

```
SELECT empName
FROM employee E
WHERE NOT EXISTS (SELECT custId
FROM customer C
WHERE C.salesRepId = E.empId
AND C.rating <> 'GOOD');
```

(a) Names of all the employees with at least one of their customers having a 'GOOD' rating.
(b) Names of all the employees with at most one of their customers having a 'GOOD' rating.
(c) Names of all the employees with none of their customers having a 'GOOD' rating.
(d) Names of all the employees with all their customers having a 'GOOD' rating.

[GATE-2014 (Set-3)]

Q.27 The relational algebra expression equivalent to the following tuple calculus expression:
$\{t \mid t \in r \wedge (t[A] = 10 \wedge t[B] = 20)\}$ is
(a) $\sigma_{(A = 10 \vee B = 20)}(r)$
(b) $\sigma_{(A = 10)}(r) \cup \sigma_{(B = 20)}(r)$
(c) $\sigma_{(A = 10)}(r) \cap \sigma_{(B = 20)}(r)$
(d) $\sigma_{(A = 10)}(r) - \sigma_{(B = 20)}(r)$

[GATE-1999]

Q.28 Which of the following relational calculus expressions is not safe?
(a) $\{t \mid \exists u \in R_1 (t[A] = u[A]) \wedge \neg \exists s \in R_2 (t[A] = s[A])\}$
(b) $\{t \mid \forall u \in R_1 (u[A] = \text{“}x\text{”} \Rightarrow \exists s \in R_2 (t[A] = s[A] \wedge s[A] = u[A]))\}$
(c) $\{t \mid \neg (t \in R_1)\}$
(d) $\{t \mid \exists u \in R_1 (t[A] = u[A]) \wedge \exists s \in R_2 (t[A] = s[A])\}$

[GATE-2001]

Q.29 With regard to the expressive power of the formal relational query languages, which of the following statements is true?
(a) Relational algebra is more powerful than relational calculus
(b) Relational algebra has the same power as relational calculus
(c) Relational algebra has the same power as safe relational calculus
(d) None of the above

[GATE-2002]

Q.30 Consider the relation employee (name, sex, supervisorName (with name as the key). supervisorName gives the name of the supervisor of the employee under consideration. What does the following Tuple Relational Calculus query produce?
{e.name | employee (e) $\wedge$ ($\forall$ x) [$\neg$ employee (x) $\vee$ x.supervisorName $\neq$ e.name $\vee$ x.sex = “male”]}
(a) Names of employees with a male supervisor
(b) Names of employees with no immediate male subordinates
(c) Names of employees with no immediate female subordinates
(d) Names of employees with a female supervisor

[GATE-2007]

Q.31 Consider a selection of the form $\sigma_{A \le 100}(r)$, where r is a relation with 1000 tuples. Assume that the attribute values for A among the tuples are uniformly distributed in the interval [0, 500]. Which one of the following options is the best estimate of the number of tuples returned by the given selection query?
(a) 50 (b) 100
(c) 150 (d) 200

[GATE IT-2007]

Q.32 Consider the following relational schema Adjacency(x, y) used to store edges of the directed graph.
Which of the following relational algebra expression results vertices set which forms loop atleast three vertices.
(a) $\pi_X\left(\text{Adjacency} \underset{\substack{Y=X_1 \\ \wedge Y_1=X}}{\bowtie} \rho_{X_1,Y_1}(\text{Adjacency})\right)$
(b) $\pi_X\left(\text{Adjacency} \underset{\substack{Y=Y_1 \\ \wedge X=X_1}}{\bowtie} \rho_{X_1,Y_1}(\text{Adjacency})\right)$
(c) $\pi_X\left(\text{Adjacency} \underset{\substack{Y=X_1 \\ Y_1 \neq X}}{\bowtie} \rho_{X_1,Y_1}(\text{Adjacency})\right)$
(d) $\pi_X\left(\text{Adjacency} \underset{\substack{Y \neq X_1 \\ Y_1=X}}{\bowtie} \rho_{X_1,Y_1}(\text{Adjacency})\right)$

Q.33 One of the following four expression of RA is not equivalent to the other three. They are based on the relation R(A, B) S(B, C)
(a) $\pi_{AB}(R \bowtie S)$
(b) $R \bowtie \pi_B(S)$
(c) $R \cap (\pi_A(R) \times \pi_B(S))$
(d) $\pi_{A, R.B}(R \times S)$

Consider the following Q.34 to Q.35:

Edges of two directed graph stored in data base relation Adj(x, y) [i.e. of there exist edge from vertex A to B then (A, B) is one record in Adj relation]

Q.34 Which of relational algebra query retries vertices who out degree atleast two?

(a) $\pi_X\left(\text{Adj} \underset{\substack{X=X_1 \\ \wedge Y \neq Y_1}}{\bowtie} \rho_{X_1,Y_1}(\text{Adj})\right)$

(b) $\pi_X\left(\text{Adj} \underset{\substack{X=X_1 \\ \wedge Y = Y_1}}{\bowtie} \rho_{X_1,Y_1}(\text{Adj})\right)$

(c) $\pi_X\left(\text{Adj} \underset{X=X_1}{\bowtie} \rho_{X_1,Y_1}(\text{Adj})\right)$

(d) $\pi_X\left(\text{Adj} \underset{\substack{X \neq X_1 \\ \wedge Y \neq Y_1}}{\bowtie} \rho_{X_1,Y_1}(\text{Adj})\right)$

Q.35 Which of the query correct representation to retrieve vertices with only degree two?

(i) SELECT X
FROM Adj
Group by X
having count (*) = 2

(ii) $\pi_X\left(\text{Adj} \underset{\substack{X=X1 \\ \wedge Y \neq Y1}}{\bowtie} \rho(\text{Adj})_{X1,\,Y1}\right)$

(a) Only (i) correct
(b) Only (ii) correct
(c) Both (i), (ii) correct
(d) Both (i), (ii) incorrect

Numerical Answer Type Questions

Q.36 The following functional dependencies hold for relations R(A, B, C) and S(B, D, E):

$B \rightarrow A$
$A \rightarrow C$

The relation R contains 200 tuples and the relation S contains 100 tuples. What is the maximum number of tuples possible in the natural join R ⋈ S?

[GATE-2010]

Q.37 Consider the following relations

R(ABC) A: Primary key with 10 tuples
S(ADE) AD: Primary key with 40 tuples
T(DFG) D: Primary key with 30 tuples

How many maximum tuples resulted by R ⋈ S ⋈ T?

Common Data for Q.38 & Q.39

Consider the following relations A, B and C:

A.

Id	Name	Age
12	Arun	60
15	Shreya	24
99	Rohit	11

B.

Id	Name	Age
15	Shreya	24
25	Hari	40
98	Rohit	20
99	Rohit	11

C.

Id	Phone	Age
10	2200	02
99	2100	01

Q.38 How many tuples does the result of the following relational algebra expression contain? Assume that the schema of A ∪ B is the same as that of A.

$(A \cup B) \bowtie_{A.Id>40 \vee C.Id<15} C$

[GATE-2012]

Q.39 How many tuples does the result of the following SQL query contain?

SELECT A.Id
FROM A
WHERE A.Age >
All (**SELECT** B.Age
FROM B
WHERE B.Name = 'Arun')

[GATE-2012]

Q.40 Consider the following relational schema:

R

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

How many records resulted by given SQL query?

```
Select *
FROM R
Where C > ALL (Select B
                FROM R
                Where A > 10)
```

Q.41 Consider a database that has the relation schema CR(StudentName, CourseName). An instance of the schema CR is as given below:

CR

Student Name	Course Name
SA	CA
SA	CB
SA	CC
SB	CB
SB	CC
SC	CA
SC	CB
SC	CC
SD	CA
SD	CB
SD	CC
SD	CD
SE	CD
SE	CA
SE	CB
SF	CA
SF	CB
SF	CC

The following query is made on the database.

$T_1 \leftarrow \pi_{\text{CourseName}}(\sigma_{\text{StudentName = 'SA'}}(\text{CR}))$

$T_2 \leftarrow \text{CR} \div T_1$

The number of rows in T_2 is _____.

[GATE-2017]

Q.42 Consider the following relational Schema and Instance

R

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

How many tuples resulted by given SQL query?

```
Select *
FROM R as T1
Where EXISTS (Select count (*)
               FROM R as T2
               Where T1 . A < T2 . A)
```

Multiple Select Questions

Q.43 Which is the subset of SQL commands used to manipulate Database Structures, including tables?

(a) Data Definition Language
(b) Data Manipulation Language
(c) Data Described Language
(d) Data Retrieval Language

Q.44 Which of the following is/are correct?

(a) Logical database design is the process of selecting the data storage and data access characteristics of the database.
(b) Physical database design is the process of selecting the data storage and data access characteristics of the database.
(c) Terminated is a state in transaction.
(d) Partially committed is not a state in transaction.

Q.45 Which of the following queries outputs the same result?

(a)
```
Select Distinct F.Name
From Faculty F
Where Not Exists
    (Select * From Class C
    Where F.Id = C.InstructorId AND
    C.Year = '2020');
```

(b) $\{F.Name \mid Faculty(F) \wedge \neg (\exists C \in Class (F.Id = C.InstructorId \wedge C.Year = '2020'))\}$

(c) $\{F.Name \mid Faculty(F) \wedge (\forall C \in Class (F.Id <> C.InstructorId \vee C.Year <> '2020'))\}$

(d) None of these

Q.46 Which of the following is/are correct regarding tuple relational calculus?

(a) $\{t \mid \neg (t \in \text{loan})\}$ may generate infinity many tuples which may not even appear in the database.

(b) A well formed query will have a single unbounded variable. All other variables will have a quantifier over them.

(c) A well formed query can have more than one free variable.

(d) If a topic variable T is bound to a relation R, then it only has values for the attributes in R. All other attribute values are null.

Try Yourself

T1. Consider the following relation

Supplies (Sid, Sname, rating)

Parts (Pid, Pname, color)

Catalog (Sid, Pid, cost)

Write the following queries in RA, TRC, SQL format

(a) Retrieve Sname's supplied at least one part

(b) Retrieve Sname's supplied some red part

(c) Retrieve Sid's supplied some red part or some green part

(d) Retrieve Sid's supplied some red part and some green part

(e) Retrieve sid's supplied only red parts

(f) Retrieve Sid's supplied every red part

(g) Retrieve Sid's supplied every red or every green part

(h) Retrieve Sid's supplied every red part but not any green part

(i) Retrieve Sid's whose rating greater than 10 and supplied every green part

(j) Retrieve red part id's which is supplied by every supplier whose rating greater than 10.

(k) Retrieve Sid's supplied most expensive part

(l) Retrieve Sid's supplied least expensive part

(m) Retrieve Sid's supplied second most expensive part

(n) Retrieve Sid's supplied most expensive for each part

(o) Retrieve Snames supplied at least two parts and whose rating greater than 10

(p) Retrieve Sid's who supplied at least two parts

(q) Retrieve Sid's who supplied only two parts

(r) Retrieve Sid's who supplied at most two parts

(s) Retrieve pairs of Sid's (Sid1, Sid2) such that supplier with Sid1 charges more than supplier with 2nd Sid (Sid2) for some part.

T2. Consider the following relation Employee (Eid, Ename, Salary, deptno, gender).

Write the SQL queries for the following questions?

(a) Retrieve Eid's who gets maximum salary

(b) Retrieve Eid's who get 2nd maximum salary

(c) Retrieve deptno for which at least two employees working on that dept

(d) Retrieve dept_no in which average salary is more than average salary in the company

(e) Retrieve dept_no in which the average salary of male employees is more than the average salary in the company.

(f) Retrieve dept_no in which the average salary of male employees is more than the average salary all female employees

(g) Retrieve dept_no in which the average salary of male employees is more than the average salary of employees in the same department

(h) Retrieve Eid's whose salary more than any employee salary of department 5

(i) Retrieve Eid's whose salary more than every employee salary of dept 5

(j) Retrieve Eid's of employees who earn more than average salary of all male employees of their dept.

T3. One of the following four expressions of relational algebra is not equivalent to the other three. They are all based on the relations R(A, B) and S(B, C). Indicate which is not equivalent to the others.

(a) $\pi_{AB}(R \bowtie S)$ (b) $R \bowtie \pi_B(S)$

(c) $R \cap (\pi_A(R) \times \pi_B(S))$ (d) $\pi_{A,R,B}(R \times S)$

T4. Consider the following relation:

family (parent, child, childDOB)

The intent is that a tuple (*p*, *c*, *d*) means that parent *p* has child *c*, who was born on date *d*. You may assume that parents do not have two children of the same name, and that their are no twins; i.e., no parent has two or more children born on the same day. Here are three queries we might ask about this data;

1. Find for each parent, the youngest child, i.e., the set of (*p*, *c*) such that *p* has child *c*, and no other child of *p* has a smaller date of birth than *c* does.
2. Find the set of great grandparents of "Amy".
3. Find all the descendants of "Mike".

Which of the above queries are expressible in relational algebra?

(a) 1 only (b) 1 and 2 only
(c) 3 only (d) 1, 2 and 3

T5. If $\cap$, $\cup$, – are given their bag interpretations, which of the following laws hold?

(a) $R \cup R = R$
(b) $R \cap (S \cup T) = (R \cap S) \cup (R \cap T)$
(c) $R \cup (S - T) = (R \cup S) - T$
(d) None of these

T6. Consider the following SQL query on the relation R(A, B) that has no NULL's.

Select rr. A, rr.B, ss.A, ss.B
From R as rr, R as ss
Where rr.A = ss.A and rr.B = ss.B

Suppose that R has *n* tuples (not necessarily all distinct). Which of the above conditions is the most restrictive correct limitation on *m*, the number of tuples (again not necessarily all distinct) in the result?

(a) $n \le m \le n^*n$ (b) $n \le m \le 2n$
(c) $0 \le m \le n$ (d) $m = n$

T7. Suppose now that R(A, B) and S(A, B) are two relations with r and s tuples, respectively (not necessarily distinct). If m is the number of (not necessarily distinct) tuples in the result of the SQL query:

R intersect S;

Then which of the following is the most restrictive, correct condition on the value of *m*?

(a) $m = \min(r, s)$
(b) $0 \le m \le r + s$
(c) $\min(r, s) \le m \le \max(r, s)$
(d) $0 \le m \le \min(r, s)$

T8. Which of the following is correct?

(a) a SQL query automatically eliminates duplicates
(b) SQL permits attribute names to be repeated in the same relation
(c) a SQL query will not work if there are no indexes on the relations
(d) None of these

T9. Suppose that two relations R(A, B) and S(A, B) have exactly the same schema. Consider the following equalities in relational algebra?

1. $R \cap S = R - (R - S)$
2. $R \cap S = S - (S - R)$
3. $R \cap S$ = R NATURAL-JOIN S
4. $R \cap S = R \times S$

Which of the equalities hold in relational algebra?

(a) 1 only
(b) 1 and 2 only
(c) 1, 2 and 3
(d) 1, 2, 3 and 4

T10. The table Arc(*x*, *y*) currently has the following tuples (not there are duplicates): (1, 2), (1, 2), (2, 3), (3, 4), (3, 4), (4, 1), (4, 1), (4, 1), (4, 2).

Compute the result of the query:

SELECT a1.x, a2.y, COUNT(*)
FROM Arc a1, Arc a2
WHERE a1.y = a2.x
GROUP BY a1.x, a2.y

Which of the following tuples is in the result?

1. (1, 3, 2)
2. (2, 4, 6)
3. (3, 1, 4)
4. (3, 1, 6)

(a) 1 and 3 (b) 1 and 2
(c) 1 and 4 (d) 2 and 3

T11. The relation R(a, b) may have duplicate tuples. Which of the following queries has a result that is guaranteed not to have duplicates, regardless of what tuples R contains?

1. SELECT a FROM R WHERE a = 1
2. SELECT MAX(b) FROM R GROUP BY a
3. SELECT a, b FROM R GROUP BY a, b
4. SELECT a FROM R WHERE a NOT IN (SELECT a FROM R)

(a) 3 and 4
(b) 1 and 2
(c) 3 only
(d) 1 and 3

T12. Let two relations r(R) and s(S). Where R and S are schemas of relation. Consider the following points.

1. The division operation $r \div s$, relation is a relation on schema R-S.
2. For $r \div s$ to be defined, $S \subseteq R$
3. For $r \div s$ to be defined, $S \supseteq R$
4. The relation $r \div s$ is a relation on schema S-R.

Which of above are true.

(a) 1 and 3 (b) 2 and 4
(c) 1 and 2 (d) 3 and 4

T13. Which of the following points are true

1. Tuple relational calculus is non-procedural.
2. Declarative data manipulation language require user to specify what data are needed and specifying how to get those data.
3. SQL is procedural DML.
4. In a procedural language, the user instruct the system to perform a sequence of operation on the database to compute the desired result.

(a) 4 only
(b) 3 and 2
(c) 1 and 4
(d) 2, 3 and 4
(e) 1 and 2

T14. What will be the number of columns and rows respectively obtained for the operation A – B, if A, B are Base union compatible and all the rows of A are common to B? Assume A has 4 columns and 20 rows; and B has 4 columns and 15 rows.

(a) 4, 0 (b) 0, 0
(c) 4, 5 (d) 8, 5

T15. Consider the following tables:

T_1			T_2	
A	B	C	C	D
1	2	5	5	8
1	3	4	9	7
2	4	5	5	6
3	6	8	4	7
3	7	8	8	6
4	8	9	9	8
4	3	5	5	7

Consider the following query on above tables:

(a) Select A
From T_1, T_2
where T_1.C == T2.C

(b) $\pi_A(T_1 \bowtie T_2)$

1. The sum of number of tuples in result of above two query is _____?
2. Let n_1 is number of tuples in first query and n_2 is second query.

Which of the following is true.

(a) $n_1 = n_2$
(b) $n_1 > n_2$
(c) $n_1 < n_2$

T16. Let the following relation schemas be given

R = (A, B, C)

S = (D, E, F)

Give an expression in the tuple relational calculus that is equivalent to each of following.

(a) $\pi_A(r)$ (b) $\sigma_{B=17}(r)$

(c) $r \times s$ (d) $\pi_{A,F}(\sigma_{C=D}(r \times s))$

T17. Consider the following database table named water_schemes:

water_schemes		
scheme_no	district_name	capacity
1	Ajmer	20
1	Bikaner	10
2	Bikaner	10
3	Bikaner	20
1	Churu	10
2	Churu	20
1	Dungargarh	10

The number of tuples returned by the following SQL query is:

```
with total (name, capacity) as
    select district_name, sum (capacity)
    from water_schemes
    group by district_name
with total_avg (capacity) as
    select avg (capacity)
    from total
    select name
    from total, total_avg
    where total.capacity ≥ total_avg. capacity
```

[GATE-2016, Ans : (2)]

■■■■

4 Transaction and Concurrency Control

Multiple Choice Questions

Q.1 Which of the following scenarios may lead to an irrecoverable error in a database system?

(a) A transaction writes a data item after it is read by an uncommitted transaction

(b) A transaction reads a data item after it is read by an uncommitted transaction

(c) A transaction reads a data item after it is written by a committed transaction

(d) A transaction reads a data item after it is written by an uncommitted transaction

[ISRO-2009]

Q.2 In the following, T1 and T2 are transactions and A is an object. Which of the following has the potential of making T2 irrecoverable?

(a) T2 writes A after T1 wrote A; T1 is uncommitted

(b) T2 reads A after T1 wrote A; T1 is uncommitted

(c) T2 write A after T1 wrote A; T1 is committed

(d) T2 reads A after T1 read A; T1 is uncommitted

Q.3 Amongst the ACID properties of a transaction, the 'Durability' property requires that the changes made to the database by a successful transaction persist

(a) Except in case of an Operating System crash

(b) Except in case of a Disk crash

(c) Except in case of a power failure

(d) Always, even if there is a failure of any kind

[GATE IT-2005]

Q.4 Consider the following log sequence of two transactions on a bank account, with initial balance 12000, that transfer 2000 to a mortgage payment and, then apply a 5% interest.

1. T1 start
2. T1 B old = 12000 new = 10000
3. T1 M old = 0 new = 2000
4. T1 commit
5. T2 start
6. T2 B old = 10000 new = 10500
7. T2 commit

Suppose the database system crashed just before log record 7 is written. When the system is restarted, which one statement is true of the recovery procedure?

(a) We must redo log record 6 set B to 10500

(b) We must undo log record 6 to set B to 10000 and then redo log records 2 and 3

(c) We need not redo log records 2 and 3 because transaction T1 has committed

(d) We can apply redo and undo operations in arbitrary order because they are idempotent

[GATE-2006]

Q.5 Consider the following schedules:

S_1: $R_1(x)\ R_2(z)\ R_1(z)\ R_3(x)\ R_3(y)\ w_1(x)\ C_1\ w_3(y)\ C_3\ R_2(y)\ w_2(z)\ w_2(y)\ C_2$

S_2: $R_1(x)\ R_2(z)\ R_1(z)\ R_3(x)\ R_3(y)\ w_1(x)\ w_3(y)\ R_2(y)\ w_2(z)\ w_2(y)\ C_1, C_2, C_3$

S_3: $R_1(x)\ R_2(z)\ R_3(z)\ R_1(z)\ R_2(y)\ R_3(y)\ w_1(x)\ C_1\ w_2(z)\ w_3(y)\ w_2(y)\ C_2, C_3$

Which of the following statement is true?

(a) S_1 cascadeless recoverable but not strict

(b) S_1, S_2 are irrecoverable schedule

(c) S_1, S_3 cascadeless recoverable and S2 not recoverable schedule

(d) S_1, S_2, S_3 all are irrecoverable schedule

Q.6 Consider the following schedule:

S_1 : $R_1(B)\ R_3(C)\ R_1(A)\ W_2(A)\ W_1(A)\ W_2(B)\ W_3(A)\ W_1(B)\ W_3(B)\ W_3(C)$

S_2 : $R_2(B)\ R_3(B)\ W_2(A)\ R_1(A)\ W_3(B)\ R_2(B)\ W_1(A)\ W_2(A)\ R_2(C)\ W_2(C)\ R_1(C)\ W_1(C)\ W_3(B)$

Which is true?

(a) S_1 and S_2 conflict serializable schedules
(b) S_1 conflict serializable but nut S_2
(c) S_2 conflict serializable but not S_1
(d) S_1 and S_2 no conflict serializable schedules

Q.7 Consider the following three schedules of transactions T_1, T_2 and T_3. [Notation: In the following NYO represents the action Y (Y for read, W for write) performed by transaction N on object O.]

(S_1) 2RA 2WA 3RC 2WB 3WA 3WC 1RA 1RB 1WA 1WB
(S_2) 3RC 2RA 2WA 2WB 3WA 1RA 1RB 1WA 1WB 3WC
(S_3) 2RA 3RC 3WA 2WA 2WB 3WC 1RA 1RB 1WA 1WB

Which of the following statements is TRUE?

(a) S_1, S_2 and S_3 are all conflict equivalent to each other
(b) No two of S_1, S_2 and S_3 are conflict equivalent to each other
(c) S_2 is conflict equivalent to S_3, but not to S_1
(d) S_1 is conflict equivalent to S_2, but not to S_3

[GATE IT-2008]

Q.8 Consider the following schedules involving two transactions. Which one of the following statements is TRUE?

S_1 : $r_1(X);\ r_1(Y);\ r_2(X);\ r_2(Y);\ w_2(Y);\ w_1(X)$

S_2 : $r_1(X);\ r_2(X);\ r_2(Y);\ w_2(Y);\ r_1(Y);\ w_1(X)$

(a) Both S_1 and S_2 are conflict serializable
(b) S_1 is conflict serializable and S_2 is not conflict serializable
(c) S_1 is not conflict serialization and S_2 is conflict serializable
(d) Both S_1 and S_2 are not conflict serializable

[GATE-2007]

Q.9 Consider the following four schedules due to three transactions (indicated by the subscript) using read and write on a data item x, denoted by $r(x)$ and $w(x)$ respectively. Which one of them is conflict serializable?

(a) $r_1(x);\ r_2(x);\ w_1(x);\ r_3(x);\ w_2(x)$
(b) $r_2(x);\ r_1(x);\ w_2(x);\ r_3(x);\ w_1(x)$
(c) $r_3(x);\ r_2(x);\ r_1(x);\ w_2(x);\ w_1(x)$
(d) $r_2(x);\ w_2(x);\ r_3(x);\ r_1(x);\ w_1(x)$

[GATE-2014 (Set-1)]

Q.10 Consider the following schedule S of transactions T_1, T_2, T_3 and T_4:

T_1	T_2	T_3	T_4
Writes(X) Commit	Reads(X)	Writes(X) Commit	
	Writes(Y) Reads(Z) Commit		Reads(X) Reads(Y) Commit

Which one of the following statements is CORRECT?

(a) S is conflict-serializable but not recoverable
(b) S is not conflict-serializable but is recoverable
(c) S is both conflict-serializable and recoverable
(d) S is neither conflict-serializable nor is it recoverable

[GATE-2014 (Set-2)]

Q.11 Consider the transactions T_1, T_2 and T_3 and the schedules S_1 and S_2 given below:

T_1 : $r_1(X);\ r_1(Z);\ w_1(X);\ w_1(Z)$

T_2 : $r_2(Y);\ r_2(Z);\ w_2(Z)$

T_3 : $r_3(Y);\ r_3(X);\ w_3(Y)$

S_1 : $r_1(X);\ r_3(Y);\ r_3(X);\ r_2(Y);\ r_2(Z);\ w_3(Y);\ w_2(Z);\ r_1(Z);\ w_1(X);\ w_1(Z)$

S_2 : $r_1(X);\ r_3(Y);\ r_2(Y);\ r_3(X);\ r_1(Z);\ r_2(Z);\ w_3(Y);\ w_1(X);\ w_2(Z);\ w_1(Z)$

Which one of the following statements about the schedules is TRUE?

(a) Only S_1 is conflict-serializable.
(b) Only S_2 is conflict-serializable.
(c) Both S_1 and S_2 are conflict-serializable.
(d) Neither S_1 nor S_2 is conflict-serializable.

[GATE-2014 (Set-3)]

Q.12 Consider the following schedule for transactions T_1, T_2 and T_3:

T_1	T_2	T_3
Read (X)		
	Read (Y)	
		Read (Y)
	Write (Y)	
Write (X)		
		Write (X)
	Read (X)	
	Write (X)	

Which one of the schedules below is the correct serialization of the above?

(a) $T_1 \rightarrow T_3 \rightarrow T_2$ (b) $T_2 \rightarrow T_1 \rightarrow T_3$

(c) $T_2 \rightarrow T_3 \rightarrow T_1$ (d) $T_3 \rightarrow T_1 \rightarrow T_2$

[GATE-2010]

Q.13 Which of the following claims correct above strict 2PL?

(i) Guaranteed conflict serializability.

(ii) Guaranteed strict recoverable.

(iii) Guaranteed deadlock avoidance.

(a) (i), (ii) only (b) (i) only

(c) (i), (ii) and (iii) (d) (ii), (iii) only

Q.14 Consider the following transactions with data items P and Q initialized to zero:

T_1 : read (P);
read (Q) ;
if P = 0 then Q: = Q + 1;
write (Q);

T_2 : read (Q);
read (P);
if Q = 0 then P: = P + 1;
write (P);

Any non-serial interleaving of T_1 and T_2 for concurrent execution leads to

(a) a serializable schedule

(b) a schedule that is not conflict serializable

(c) a conflict serializable schedule

(d) a schedule for which a precedence graph cannot be drawn

[GATE-2012]

Q.15 For the schedule given below, which of the following is correct?

T_1	T_2
R(A)	
	R(B)
W(A)	
	R(A)
	W(A)
	W(B)
R(B)	
W(B)	

(a) This schedule is serializable and can occur in a scheme using 2PL protocol.

(b) This schedule is serializable but cannot occur in a scheme using 2PL protocol.

(c) This schedule is not serializable but can occur in a scheme using 2PL protocol.

(d) This schedule is not serializable and cannot occur in a scheme using 2PL protocol.

[GATE-1999]

Q.16 Consider the following locking protocol:

(i) Transaction must lock all data items before begins any read and write operation.

(ii) Transaction executes all read and write operation until transaction commit/rollback.

(iii) Transaction unlocks all data items.

Which of the following is true above given protocol?

(a) Guaranteed no deadlock and guaranteed strict recoverability may not serializable

(b) Guaranteed no deadlock and guaranteed no starvation

(c) Guaranteed no deadlock and guaranteed serializability and strict recoverability

(d) Guaranteed serializable and no deadlocks mays not strict recoverable

Q.17 Consider the following two transactions: T_1 and T_2.

T_1:	T_2:
read (A);	read (B);
read (B);	read (B);
If A = 0 then B ← B+1	if B ≠ 0 then A ← A – 1;
write (B);	write (A);

Which of the following schemes, using shared and exclusive locks, satisfy the requirements for strict two locking for the above transactions?

(a)

S_1 :	S_2 :
lock S(A);	lock S (B);
read (A);	read (B);
lock S(B);	lock S(A);
read (B);	read (A);
if A = 0	if B ≠ 0
then B←B+1;	then A←A−1;
write (B);	write (A);
commit;	commit;
unlock(A);	unlock (B);
unlock (B);	unlock (A);

(b)

S_1 :	S_2 :
lock X(A);	lock X (B);
read (A);	read (B);
lock X(B);	lock X(A);
read (B);	read (A);
if A = 0	if B ≠ 0
then B←B+1;	then A→A−1;
write (B);	write (A);
unlock(A);	unlock (A);
commit	commit
unlock (B);	unlock (B);

(c)

S_1 :	S_2 :
lock S(A);	lock S (B);
read (A);	read (B);
lock X(B);	lock X(A);
read (B);	read (A);
if A = 0	if B ≠ 0
then B←B+1;	then A←A−1;
write (B);	write (A);
unlock(A);	unlock (B);
commit	commit;
unlock (B);	unlock (A);

(d)

S_1 :	S_2 :
lock S(A);	lock S(B);
read (A);	read (B);
lock X(B);	lock X(A);
read (B);	read (A);
if A = 0	if B ≠ 0
then B←B+1;	then A←A−1;
write (B);	write (A);
unlock(A);	unlock(A);
unlock (B);	unlock (B);
Commit;	Commit;

[GATE IT-2007]

Q.18 In a database system, unique timestamps are assigned to each transaction using Lamport's logical clock. Let $TS(T_1)$ and $TS(T_2)$ be the timestamps of transactions T_1 and T_2 respectively. Besides, T_1 holds a lock on the resource R and T_1 has requested a conflicting lock on the same resource R. The following algorithm is used to prevent deadlocks in the database system assuming that a killed transaction is restarted with the same timestamp.

if $TS(T_2) < TS(T_1)$ then
  T_1 is killed
else T_2 waits

Assume any transaction that is not killed terminates eventually. Which of the following is TRUE about the database system that uses the above algorithm to prevent deadlocks?

(a) The database system is both deadlock-free and starvation-free.
(b) The database system is deadlock-free, but not starvation-free.
(c) The database system is starvation-free, but not deadlock-free.
(d) The database system is neither deadlock-free nor starvation-free.

[GATE-2017]

Q.19 Consider the following schedule:

S : $r_2(A)$, $w_1(B)$, $w_1(C)$, $r_3(B)$, $r_2(B)$, $r_1(A)$, c_1 $r_2(C)$, c_2 $w_3(A)$ c_3

How many given statements true about schedule(S)?

(i) Schedule(S) is conflict serializable schedule.
(ii) Schedule(S) is allowed by 2PL.
(iii) Schedule(S) is strict recoverable schedule.
(iv) Schedule(S) is allowed by strict 2PL.

(a) 1 (b) 2
(c) 3 (d) 4

Q.20 Consider the following schedule:

S : $w_1(A)$, $w_1(B)$, $r_2(A)$, $w_2(B)$, $r_3(A)$, $w_3(B)$

How many schedules are conflict equivalent to given schedule(S)?

(a) 10 (b) 15
(c) 8 (d) 1

Consider Data for Q.21 and Q.22:
Consider the following schedule
S : $r_1(A)$, $r_3(D)$, $w_1(B)$, $r_2(B)$, $r_4(B)$, $w_2(C)$, $r_5(C)$, $w_4(E)$, $r_5(E)$, $w_5(B)$

Q.21 How many serial schedules conflict equal to schedules (S)?
(a) 10 (b) 15
(c) 8 (d) 12

Q.22 How many serial schedules view equal to schedule (S)
(a) 10 (b) 15
(c) 8 (d) 12

Numerical Answer Type Questions

Q.23 Two transactions T_1 and T_2 are given as
T_1 : $r_1(X)$ $w_1(X)$ $r_1(Y)$ $w_1(Y)$
T_2 : $r_2(Y)$ $w_2(Y)$ $r_2(Z)$ $w_2(Z)$
where $r_i(V)$ denotes a read operation by transaction T_i on a variable V and $w_i(V)$ denotes a write operation by transaction T_i on a variable V. The total number of conflict serializable schedules that can be formed by T_1 and T_2 is _____.
[GATE-2017]

Q.24 Consider the following transactions
T_1 : $r_1(A)$ $w_1(A)$ $r_1(B)$ $w_1(B)$
T_2 : $r_2(A)$ $r_2(B)$
T_3 : $w_3(A)$ $w_3(B)$
How many concurrent schedules between T_1, T_2 and T_3 transactions _____.

Q.25 How many view equivalent serial schedules are possible for the given schedule
S : $w_1(A)$ $r_2(A)$ $w_3(A)$ $r_4(A)$ $w_5(A)$ $r_6(A)$

Q.26 Consider the following transactions:
T_1 : $r_1(A)$ $w_1(A)$ $r_1(B)$ $w_1(B)$
T_2 : $w_2(B)$ $w_2(A)$
How many non serial schedules are serializable?

Q.27 Consider the following transactions
T_1 : $r_1(A)$ $r_1(B)$ $w_1(B)$
T_2 : $r_2(A)$ $r_2(B)$ $w_2(B)$
How many non serial schedules between T_1 and T_2 are serializable?

Q.28 Consider the following transactions
T_1 : $r_1(A)$ $r_1(B)$ $w_1(B)$
T_2 : $r_2(B)$ $r_2(A)$ $w_2(B)$
How many non serial schedules between T_1 and T_2 are serializable?

Q.29 Consider the following transactions
T_1 : $r_1(A)$ $w_1(A)$ $r_1(B)$ $w_1(B)$
T_2 : $r_2(B)$ $w_2(B)$ $r_2(A)$ $w_2(A)$
(*i*) How many schedules serializable as $T_1 \rightarrow T_2$
(*ii*) How many schedules serializable as $T_2 \rightarrow T_1$

Q.30 Consider the following transactions
T_1 : $r_1(A)$ $w_1(A)$ $r_1(B)$ $w_1(B)$
T_2 : $r_2(A)$ $w_2(A)$ $r_2(B)$ $w_2(B)$
(*i*) How many schedules serializable as $T_1 \rightarrow T_2$
(*ii*) How many schedules serializable as $T_2 \rightarrow T_1$

Q.31 Consider the following schedule
S : $r_1(A)$ $w_1(B)$ $r_2(A)$ $w_2(B)$ $r_3(A)$ $w_3(B)$
(*i*) How many schedules conflict equal to given schedule S?
(*ii*) How many schedules view equal to given schedule S?

Multiple Select Questions

Q.32 Which of the following is/are correct?
(a) Isolation means that the data used during the execution of a transaction cannot be used by a second transaction until the first one is completed.
(b) Either all operations of the transaction are reflected properly in the database or none, is property of Atomicity.

(c) The property of a transaction that persists all the crashes is Durability
(d) Consistency states that only valid data will be written to the database.

Q.33 Which of the following is/are correct?
(a) Two phase lock is a procedure for acquiring the necessary locks for a transaction where all necessary locks are acquired before any are released.
(b) When transaction T_i requests a data item currently held by T_j, T_i is allowed to wait only if it has a timestamp smaller than that of T_j (that is, T_i is older than T_j). Otherwise, T_i is rolled back (dies). This is Wait-die system.
(c) When transaction T_i requests a data item currently held by T_j, T_i is allowed to wait only if it has a timestamp smaller than that of T_j (that is, T_i is older than T_j). Otherwise, T_i is rolled back (dies). This is Wound-wait system.
(d) When transaction T_i requests a data item currently held by T_j, T_i is allowed to wait only if it has a timestamp larger than that of T_j (that is, T_i is younger than T_j). Otherwise, T_j is rolled back (T_j is wounded by T_i). This is Wound-wait system.

Q.34 Which of the following is/are correct regarding deadlock in transaction?
(a) In deadlock prevention, a transaction requesting a new lock is aborted if there is the possibility that a deadlock can occur.
(b) In deadlock prevention, if the transaction is aborted, all the changes made by this transaction are rolled back, and all locks obtained by the transaction are released.
(c) In deadlock avoidance, the transaction must obtain all the locks it needs before it can be executed.
(d) In deadlock detection, the DBMS periodically tests the database for deadlocks.

Q.35 Which of the following schedules is/are conflict serializable?
(a) $r_1(x); r_3(x); w_1(x); r_2(x); w_3(x)$
(b) $r_1(x); r_3(x); w_3(x); w_1(x); r_2(x)$
(c) $r_3(x); r_2(x); w_3(x); r_1(x); w_1(x)$
(d) $r_3(x); r_2(x); r_1(x); w_3(x); w_1(x)$

Try Yourself

T1. S_1: $r_1(A)\ r_2(A)\ r_3(A)\ w_1(B)\ w_2(B)\ w_3(B)$

S_2: $r_1(A)\ r_2(A)\ r_3(A)\ r_4(A)\ w_1(B)\ w_2(B)\ w_3(B)\ w_4(B)$

S_3: $r_1(A)\ r_3(D)\ w_1(B)\ r_2(B)\ w_3(B)\ r_4(B)\ w_2(C)\ r_5(C)\ w_4(E)\ r_5(E)\ w_5(B)$

S_4: $w_1(A)\ r_2(A)\ w_3(A)\ r_4(A)\ w_5(A)\ r_6(A)$

S_5: $r_2(A)\ r_1(A)\ w_1(C)\ r_3(C)\ w_1(B)\ r_4(B)\ w_3(A)\ r_4(C)\ w_2(D)\ r_2(B)\ w_4(A)\ w_4(B)$

For the above schedules find all view equivalent serials orders.

T2. Consider the following schedules S_1, S_2, S_3 below. Determine whether each schedule is strict, cascadeless, recoverable or non recoverable.

S_1: $r_1(x)\ r_2(z)\ r_1(z)\ r_3(x)\ r_3(y)\ w_1(x)\ C_1\ w_3(y)\ C_3\ r_2(y)\ w_2(z)\ w_2(y)\ C_2$

S_2: $r_1(x)\ r_2(z)\ r_1(z)\ r_3(x)\ r_3(y)\ w_1(x)\ w_3(y)\ r_2(y)\ w_2(z)\ w_2(y)\ C_1\ C_2\ C_3$

S_3: $r_1(x)\ r_2(z)\ r_3(x)\ r_1(z)\ r_2(y)\ r_3(y)\ w_1(x)\ C_1\ w_2(z)\ w_3(y)\ w_2(y)\ C_3\ C_2$

T3.
1. $R_1(x)\ R_2(x)\ w_1(x)\ w_2(x)$
2. $w_1(x)\ R_2(y)\ R_1(y)\ R_2(x)$
3. $R_1(x)\ R_2(y)\ w_3(x)\ R_2(x)\ R_1(y)$
4. $R_1(x)\ R_2(y)\ w_1(x)\ R_2(y)\ w_3(y)\ w_1(x)\ R_2(y)$
5. $R_1(x)\ w_2(x)\ w_1(x)$ Abort 2; commit 1
6. $r_1(x)\ w_2(x)\ w_1(x)\ C_2, C_1$
7. $w_1(x)\ R_2(x)\ w_1(x)\ R_2, C_1$
8. $w_1(x)\ R_2(x)\ w_1(x)\ C_2, C_1$

Consider the following classes of schedules identify

- Conflict serializable
- View serializable
- Recoverable
- Avoid cascading aborts
- Strict recoverable

T4. Consider the following schedules:

(a) $r_1(A)\, r_2(B)\, r_3(C)\, w_1(B)\, w_2(C)\, w_3(D)$

(b) $r_1(A)\, r_2(B)\, r_3(C)\, w_1(B)\, w_2(C)\, w_3(A)$

(c) $r_1(A)\, r_2(B)\, r_3(C)\, r_1(B)\, r_2(C)\, r_3(D)\, w_1(C)\, w_2(D)\, w_3(E)$

(d) $r_1(A)\, r_2(B)\, r_3(C)\, r_1(B)\, r_2(C)\, r_3(D)\, w_1(A)\, w_2(A)\, w_3(C)$

(e) $r_1(A)\, r_2(B)\, r_3(C)\, r_1(B)\, r_2(C)\, r_3(A)\, w_1(A)\, w_2(B)\, w_3(C)$

Answer the following questions for the above schedules.

(i) Find all conflict equal serial schedules?

(ii) Find all view equal serial schedules?

(iii) Find the schedules which are possible to execute by basic 2PL? (Use lock upgrading if required)

(iv) Find the schedules which are possible to execute by strict 2PL protocol? Consider commit operation as immediate last operations of each trans.

Example: $r_1(A)$, $r_2(B)$, $r_3(C)$, $w_1(B)$, $commit_1$, $w_2(C)$ $commit_2$, $w_3(D)$ $commit_3$

T5. Isolation of the transactions is ensured by

(a) Transaction management
(b) Application programmer
(c) Concurrency control
(d) Recovery management

T6. Ensuring consistency for an individual transaction is the responsibility of _____.

(a) Concurrency control component
(b) Transaction management
(c) Application programmer
(d) Recovery management

T7. Consider the following point about serializability. Which are false

1. Testing conflict serializability is polynomial problem i.e., it can be done in polynomial time.
2. Testing view serializability is NP component problem.
3. Every view serializable is also conflict serializable.
4. View serializability is necessary but not sufficient for serializability of schedule.

(a) 1 and 2
(b) 3 and 4
(c) 2, 3 and 4
(d) 1, 3 and 4

[Ans: (b)]

T8. Which of the following must be idempotent

(a) Commit (b) Redo
(c) Write (d) Undo

T9. Which of the following is false

1. Wait-die scheme for deadlock prevention is non-preemptive and wound-wait is preemptive.
2. Wait-die is preemptive and wound-wait is non-preemptive.
3. In wait-die scheme there is no deadlock and avoid starvation.
4. For wound-wait scheme there is possibility of starvation.

(a) 2 and 4 (b) 1 and 3
(c) 2 and 3 (d) 1 and 4

T10. Which of the following is right about 2PL

1. If schedule is conflict serializable then it is allowed by 2PL.
2. 2PL may lead to deadlock.
3. 2PL allowed schedule are free from cascading roll back and lost update problem
4. Starvation may happen.

(a) 1 and 4 (b) 2 and 3
(c) 2 and 4 (d) 1 and 3

[Ans: (c)]

T11. Which of the following are true about basic time stamp ordering protocol.

1. Free From Deadlock
2. Free From Starvation
3. Not Free From Cascading Rollback problem and Irrecoverable schedules
4. If schedule is conflict serializable, this condition alone is sufficient to tell that schedule is allowed by BTSO protocol.

(a) 1 and 3
(b) 1 and 4
(c) 2 and 4
(d) 2, 3 and 4

T12. Consider the following database schedule with two transactions, T_1 and T_2.

$S = r_2(X); r_1(X); r_2(Y); w_1(X); r_1(Y); w_2(X); a_1; a_2$

where $r_i(Z)$ denotes a read operation by transaction T_i on a variable Z, $w_i(Z)$ denotes a write operation by T_i on a variable Z and a_i denotes an abort by transaction T_i.

Which one of the following statements about the above schedule is **TRUE**?

(a) S is non-recoverable
(b) S is recoverable, but has a cascading abort
(c) S does not have a cascading abort
(d) S is strict

[GATE-2016, Ans: (c)]

■■■■

5 File Structure and Indexing

Multiple Choice Questions

Q.1 Choose the correct statements.
(a) For fixed length records unspanned organization is preferred.
(b) For variable length records unspanned organization is preferred.
(c) For fixed length records spanned organization is preferred.
(d) None of the above

Q.2 Order P is maximum child pointers per B tree node.
How many minimum levels of index required to store 300 distinct keys in B tree with order P is 5?
(a) 4 (b) 3
(c) 5 (d) 2

Q.3 Suppose blocks hold either three records, or Ten (Key, pointer) pairs. As a function of n, the number of records, How many blocks do we need to hold a data file and a dense index.
(a) $\frac{13n}{30}$ (b) $\frac{11n}{30}$
(c) $\frac{10n}{30}$ (d) $\frac{n}{3}$

Q.4 Suppose blocks hold either three records, or Ten (Key, pointer) pairs. As a function of n, the number of records, for dense index. How many levels of index as is appropriate until the final level of index has only one block?
(a) $\frac{n}{3^n}$ (b) $\text{Log}_3 n$
(c) $\frac{n}{4^n}$ (d) $\text{Log}_{10} n$

Q.5 Consider a file of 16384 records. Each record is 32 bytes long and its key fields of size 6 bytes. The file is ordered on a non-key field, and the file organization is unspanned. The file is stored in a file system with block size 1024 bytes, and the size of a block pointer is 10 bytes. If the secondary index is built on the key field of the file, and a multi-level index scheme is used to store the secondary index, the number of first-level and second-level blocks in the multi-level index are respectively
(a) 8 and 0 (b) 128 and 6
(c) 256 and 4 (d) 512 and 5
[GATE-2008]

Q.6 A clustering index is defined on the fields which are of type
(a) Non-key and ordering
(b) Non-key and non-ordering
(c) Key and ordering
(d) Key and non-ordering
[GATE-2008]

Q.7 For given database file, if records physically ordered based on candidate key field (x) and field x used for index. The index said to be
(a) Primary index
(b) Clustering index
(c) Secondary index on key
(d) Secondary index on non-key

Q.8 For given database file, if records physically ordered based on candidate key field x and non key field (y) used for index. Then index said to be
(a) Primary index
(b) Clustering index
(c) Secondary index on key
(d) Secondary index on non-key

Q.9 For given database file, if records physically ordered based on non key field (x) and index build based on candidate key field (y). Then index said to be
(a) Primary index
(b) Clustering index
(c) Secondary index on key
(d) Secondary index on non-key

Q.10 For given DB file, if records physically ordered based on non key field (x) and index build on same non key field (x). Then index said to be
(a) Primary index
(b) Clustering index
(c) Secondary index on key
(d) Secondary index on non-key

Q.11 Consider file consist 65, 536 records. Each record 32 byte long and its search key field size 6 bytes. Disk block size 1024 bytes. Size of pointer 12 bytes.
How many index blocks, levels of index, I/O cost are required in worst case if
(a) Index build on the **key field (x)** of the file and records physically ordered based on **non key field (y)**.
(b) Index build on the **key field (x)** of the file and records physically ordered based on same **key field (x)**.
(c) Index build on the **non key field (x)** of the file and records physically ordered based on same **non key field (x)** and every 16 records whose non key field value same.
(d) Index build on the **non key field (x)** if the file and records physically ordered based **on key field (y)**.

Q.12 Order P is between 2 to P child pointers can store in root node and between [P] to P child pointers can store in other 2 nodes of B tree index. How many maximum levels of index required to store 300 distinct keys in order (P) 5 B tree index?
(a) 4 (b) 5
(c) 6 (d) 7

Q.13 There are five records in a database.

Name	Age	Occupation	Category
Rama	27	CON	A
Abdul	22	ENG	A
Jeniffer	28	DOC	B
Maya	32	SER	D
Dev	24	MUS	C

There is an index file associated with this and it contains the values 1, 3, 2, 5 and 4. Which one of the fields is the index built from?
(a) Age (b) Name
(c) Occupation (d) Category
[GATE-1998]

Q.14 Which statement is false?
(a) The leaf nodes of the B^+ tree are usually linked together to provide ordered access on the search field to the records.
(b) Most implementations of dynamic multilevel index use a variation of the B tree data structure called as B^+ tree.
(c) For a B^+ tree constructed on a key, the pointers in internal nodes are tree pointers to blocks that are tree nodes.
(d) In B^+ tree, data pointers are stored in the leaf nodes of the tree and the structure of leaf nodes is same as that of the internal nodes.

Q.15 Which one of the following is a key factor for preferring B^+ trees to binary search trees for index database relations?
(a) Database relations have a large number of record
(b) Database relations are sorted on the primary key
(c) B^+ trees require less memory than binary search trees
(d) Data transfer from disks is in blocks
[GATE-2005]

Q.16 Which of the following is correct?
(a) B-trees are for storing data on disk and B^+ trees are for main memory.
(b) Range queries are faster on B^+ trees.
(c) B-trees are for primary indexes and B^+ trees are for secondary indexes.
(d) The height of a B^+ tree is independent of the number of records.

[GATE-1999]

Q.17 B^+ trees are preferred to binary trees in databases because
(a) Disk capacities are greater than memory capacities
(b) Disk access is much slower than memory access
(c) Disk data transfer rates are much less than memory data transfer rates
(d) Disks are more reliable than memory

[GATE-2000]

Common Data for Q.18 & Q.19

A DB table T1 has 2000 records and occupies 80 disk block. Another table T2 has 400 records and occupies 20 disk block. These two tables have to be jointed as per a specified join condition that needs to be evaluated for every pair of records from those two tables. The memory buffer space available can hold exactly one block of record file for T1 and one block of records for T2 simultaneously at any point of time.

Q.18 If nested loop join algorithm is employed to perform the join, with the most appropriate choice of table to be used in order loop. The number blocks access required for reading the data are
(a) 8,00,000 (b) 40,080
(c) 32,020 (d) 100

Q.19 If instead of nested loop join, block nested loop join is used, again with the most appropriate choice of table in the outer loop the reduction in number block accesses required for reading the data will be
(a) 0 (b) 30,400
(c) 38,400 (d) 798,400

Q.20 Consider a join (relation algebra) between relations $r(R)$ and $s(S)$ using the nested loop method. There are 3 buffers each of size equal to disk block size, out of which one buffer is reserved for intermediate results.

Assuming size($r(R)$) < size($s(S)$), the join will have fewer number of disk block accesses if
(a) relation $r(R)$ is in the outer loop.
(b) relation $s(S)$ is in the outer loop.
(c) join selection factor between $r(R)$ and $s(S)$ is more than 0.5.
(d) join selection factor between $r(R)$ and $s(S)$ is less than 0.5.

[GATE-2014 (Set-2)]

Q.21 Which of the following statement false for B tree and B^+ tree index.
(a) B^+ tree index faster for range queries compare to B tree index.
(b) If disk block allocated for B^+ tree index and same size disk block allocated for B tree index then number of index blocks and I/O cost of B^+ tree index loss than or equal to B tree index for given distinct keys.
(c) If disk block allocated for B^+ tree index and same size disk block allocated for B tree index. Then B tree index access cost loss than or equal to B^+ tree index for given distinct keys.
(d) If number of keys that can store in B tree and B^+ tree index is same then I/O cost of B tree index less than equal to I/O cost of B^+ tree index for random access of some key from set of distinct keys.

Q.22 Which of the statement correct if maximum possible key that can store in B tree node equal to that of B^+ tree node?
(a) B tree index nodes less than or equal B^+ tree index nodes for n keys.
(b) B^+ tree index nodes less than or equal B tree index nodes for n keys.
(c) I/O cost of B^+ tree index less than or equal I/O cost of B tree index for n keys.
(d) Levels of B^+ tree index less than or equal to levels of B tree index for n keys.

Numerical Answer Type Questions

Q.23 Suppose we have a block-addressable disk drive. With such block-organized disk non-data overhead of subblocks and interblock gaps have to be accounted for. There are 40000 bytes per track and the amount of space taken up by subblocks and interblocks gaps equivalent to 250 bytes per block. A file contains records and record size is 200 bytes to be stored on the disk. If a total of 32 blocks can be stored per track then what is the blocking factor? The term "blocking factor" is used to indicate the number of records that are to the stored in each block in a file. A block is organised to hold an integral number of logical records.

Q.24 Given a system using unspanned blocking and 100-byte blocks. A file contains records of 20, 50, 35, 70, 40, 20.
What percentage of space will be wasted in the blocks allocated for the file?

Q.25 Consider a table T in a relational database with a key field K. AB tree of order p is used as an access structure on K, where p denotes the maximum number of tree pointers in a B tree node. Assume that K is 10 bytes long; disk block size is 512 bytes; each data pointer P_D is 8 bytes long and each block pointer P_B is 5 bytes long. In order for each B tree node to fit in a single disk block, the maximum value of p is _____.

[GATE IT-2004]

Q.26 A B^+ tree index is to be built on the Name attribute of the relation STUDENT. Assume that all student names are of length 8 bytes, disk block are size 512 bytes, and index pointer are of size 4 bytes. Given this scenario, what would be the best choice of the degree (i.e. the number of pointers per node) of the B^+ tree?

[GATE-2002]

Q.27 What will be the order (p) of a B^+ tree with a database of 5,00,000 records of 200 bytes each and the search key is 15 bytes? Assume the tree and data pointers are 5 bytes each and the index node (data block size) is 1024 bytes.

Q.28 The order of an internal node in a B^+ tree index is the maximum number of children it can have. Suppose that a child pointer takes 6 bytes, the search field value takes 14 bytes, and the block size is 512 bytes. What is the order of the internal node?

[GATE-2004]

Q.29 The order of a leaf node in a B^+ tree is the maximum number of (value, data record pointer) pairs it can hold. Given that the block size is 1 K bytes, data record pointers is 7 bytes long, the value field is 9 bytes long and a block pointer is 6 bytes long, what is the order of the leaf node?

[GATE-2007]

Q.30 In a database file structure, the search key field is 9 bytes long, the block size is 512 bytes, a record pointer is 7 bytes and a block pointer is 6 bytes. The largest possible order of a non-leaf node in a B^+ tree implementing this file structure is _____?

[GATE IT-2006]

Q.31 Assume block size 4096 bytes, size of key is 4 bytes. Size of pointer be 8 bytes. How many keys are possible per blocks for B^+-Tree organization?

Q.32 Find minimum and maximum nodes and keys in B/B+tree with order P = 7 and level 5
Order P : For root node key between 1 to P other nodes keys between $\lceil P/2 \rceil$ to P.

Q.33 Find minimum and maximum nodes and keys in B/B+tree with order P = 7 and level 5.
Order P : For root node between 2 to 2P child pointers.
Other internal nodes between P to 2P child pointers.
For leaf nodes between (P–1) to (2P–1) keys.

Q.34 Assume that you have built a dense primary B+-Tree indeed on a file containing 20,000 records. The key field for this B+-Tree indeed is a 40 byte string and it is a candidate key. Pointers are at 10 bytes. The size of one disk page is 1000 bytes. The index was built in a bottom-up fashion using the bulk-loading algorithm and nodes were filled up as much as possible. How many levels of indexing is required?

Q.35 A B+-Tree of order d is a tree in which each internal node has between d and $2d$ keys values.

The root has between 1 and $2d$ keys values what is the maximum number of internal nodes in a B+ -Tree of order 4 with 52 leaves.

Q.36 The following key values are inserted into a B+ -Tree in which order of the internal nodes is 3 and that of the leaf nodes is 2, in the sequence given below. The order of internal nodes is the maximum number of tree pointers in each node, and the order of leaf nodes is the maximum number of data items that can be stored in it. The B+ - Tree is initially empty.

10, 3, 6, 8, 4, 2, 1

The maximum number of times leaf nodes would get split up as a result of these insertions is

[GATE-2009]

Q.37 Consider a B+ - Tree in which the maximum number of keys in a node is 5. What is the minimum number of keys in any non-root node?

[GATE-2010]

Q.38 Assume DB table T_1 has 2000 records and occupies 80 disk blocks. Another table T_2 has 400 records and 20 disk blocks.

How many min number of main memory blocks required to join T_1 and T_2 using Nested loop join algorithm so that T_1 and T_2 should avoid repeated access from secondary memory to main memory?

Multiple Select Questions

Q.39 Consider a B-tree with degree d and the number of children C of any internal node except the root is such that $d \le C \le 2d - 1$. Which of the following are correct?

(a) The minimum key possible at height h is $2 * d^h * (d - 1)$.

(b) The minimum key possible at height h is $2 * d^{h-1} * (d - 1)$.

(c) The maximum key possible at height h is $(2d - 1)^h * (2d - 2)$.

(d) The maximum key possible at height h is $2d^h * (2d - 2)$.

Q.40 Which of the following is a key factor for preferring B^+ trees to binary search trees for indexing database relations?

(a) Database relations have a large number of records.

(b) Database relations are sorted on the primary key.

(c) B^+ trees require more memory than binary search trees.

(d) Data transfer from disks is in blocks.

Q.41 Consider a file of 16384 records where each records are 32 bytes long and its key field is of 6 bytes. The file is ordered on a non-key field and the file organization is unspanned. The file is stored in a file system with block size 1024 bytes and size of a block pointer is 10 bytes. If the secondary index is built on the key field of the file and a multi-level index scheme is used to store the secondary index. Which of the following is correct?

(a) The number of blocks in the first level will be 512.

(b) The number of blocks in the first level will be 256.

(c) The number of blocks in the second level in the multi-level index will be 4.

(d) The number of blocks in the second level is 9.

Q.42 A file consists of records of size 64 bytes each including keyfied of size 14 bytes. An address of a disk block takes 2 bytes. Assume the disk block size is 512 bytes and there are 16 K records then which of the following is true?
(a) Blocking factor of record file is 8.
(b) Blocking factor of index file is 24.
(c) Number of blocks in the first level is 64.
(d) Number of blocks in the second level index is 10.

Q.43 A database table T_1 has 2000 records and occupies 80 disk blocks. Another table T_2 has 400 records and occupies 20 disk blocks. These two tables have to be jointed as per a specified join condition that needs to be evaluated for every pair of records from these two tables. The memory buffers space available can hold exactly one block of records for T_1 and one block of records for T_2 simultaneously at any point in time.
No index is available on either table. Which of the following is correct?
(a) T_2 as the outer table in nested loop join algorithm for reading the data
(b) The number of block access will be 32020 in nested loop join algorithm
(c) In block nested loop join number of block access is 1600.
(d) In nested loop join number of block access is 28080.

Q.44 Which of the following is/are correct?
(a) In ordered indices the file containing the records is sequentially ordered, a Clustered index is an index whose search key also defines the sequential order of the file.
(b) In ordered indices the file containing the records is sequentially ordered, a Non clustered is an index whose search key also defines the sequential order of the file.
(c) Indices whose search key specifies an order different from the sequential order of the file are called Non-clustered indices.
(d) Indices whose search key specifies an order different from the sequential order of the file are called Secondary indices.

Q.45 Which of the following is/are correct?
(a) In a Dense clustering index, the index record contains the search-key value and a pointer to the first data record with that search-key value and the rest of the records will be in the sequential pointers.
(b) In a Sparse clustering index, the index record contains the search-key value and a pointer to the first data record with that search-key value and the rest of the records will be in the sequential pointers.
(c) In Sparse the indices values are larger, index is created for these values of the index. This is called multilevel index.
(d) In Sparse the indices values are larger, index is created for these values of the index. This is called Sequential index.

T1. Consider file consist of 10,000 record. Block size 1024 bytes, record size 100 bytes. Search key size 9 bytes, pointer 7 bytes.
(a) How many 1st level index blocks using dense index.
(b) How many 1st level index blocks using of sparse index.
(c) How many levels of index required if 1st level dense index.
(d) How many levels of index required if 1st level sparse index.

T2. Consider a file consist 30,000 fixed length records of size 100 bytes. Each disk block size is 1024 bytes, block pointer size 6 bytes. Search key size 9 bytes.
(a) How many 1st level index blocks using dense index?

(b) How many 1st level index blocks using sparse index.
(c) How many levels of index required if 1st level uses dense index.
(d) How many levels of index required if 1st level uses sparse index.

T3. DB file consist 1250 records. Block can hold either 3 records or (10 keys, 11 pointers)
(a) How many minimum index blocks and minimum levels of index required. Dense B^+ tree index for above DB file.
(b) How many maximum index blocks and maximum levels if index required for dense B^+ tree. Index for above DB file.
(c) How many min index blocks and min levels of index required for sparse B^+ tree index for above DB file.
(d) How many max index blocks and max levels of index required for sparse B^+ tree. Index for above DB file.

T4. With respect to the B^+ tree index method select the true statements
(a) Records are physically stored in primary key order
(b) B^+ tree uses a hashing algorithm
(c) The index tree may or may not become unbalanced as a result of updates
(d) None of the above

T5. A B-tree used as an index for a large database table has four levels including the root node. If a new key 18 inserted in this index, then the maximum number of nodes that could be newly created in the process are
(a) 3 (b) 4
(c) 5 (d) 2

T6. A data dictionary doesn't provide information about
(a) where data is located
(b) the size of the disk storage
(c) who owns the data
(d) how data is used

T7. B^+ Trees are considered BALANCED because
(a) the lengths of the paths from the root to all leaf nodes are all equal.
(b) the lengths of the paths from the root to all leaf nodes differ from each other by atmost 1.
(c) the number of children of any two non-leaf sibling nodes differ by at most 1.
(d) the number of records in any two leaf nodes differ by at most 1.

[GATE-2016, Ans : (a)]

T8. Find the minimum number of elements (keys) that need to be inserted into a B^+ tree with the order of internal node 3[maximum child pointers per node] and leaf node 2[maximum keys per node], to make it to reach 3 levels?

[Ans : (7)]

T9. Suppose size of block 1000 bytes search key of 12 bytes, pointer size 8 bytes. How many maximum records in DB file which can index by dense B^+ tree of 2 levels?
(a) 1200 (b) 1250
(c) 15000 (d) 2499

[Ans : (d)]

T10. DB file consist 50000 records with record size 100 byte block size 512 bytes. If sparse B^+ tree index build over given DB file with search key size 20 bytes both block pointer and record pointer sizes 12 bytes each.
How many max index blocks required if node order P is defined as between $\left\lceil \frac{P}{2} \right\rceil$ to P pointers per node?
(a) 1635 (b) 1631
(c) 716 (d) 713

[Ans : (b)]

■■■■