GATE 2020
PSUs 2020

MADE EASY Publications

Revised & Updated Edition

POSTAL
STUDY PACKAGE

COMPUTER SCIENCE & IT
Database Management System

Objective Practice Sets

POSTAL Study Package 2020

Computer Science & IT

Objective Practice Sets

Database Management System

Contents

MADE EASY Publications

1 CHAPTER
The Relational Model

Q.1 Which one of the following NULL values can not be used in the relational model?
(a) To fill a column in tuple when that column does not really exist for that particular tuple.
(b) To leave columns in a tuple marked as 'unknown' when the actual value is unknown.
(c) To allow duplicate tuples in the table by filling the primary key column(s) with NULL.
(d) To opt a tuple out of enforcement of a foreign key.

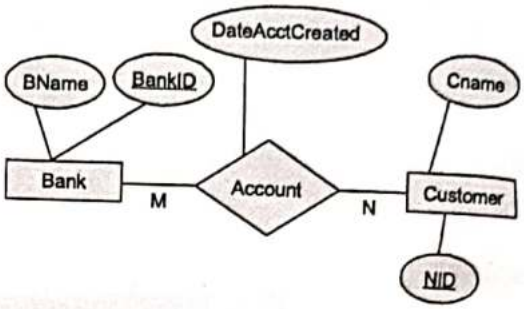
Q.2 A weak entity _____.
(a) is an entity with no attributes beside its key.
(b) inherits part of its key from the 'parent' entities to which it is related.
(c) is an entity with no key.
(d) None of these.

Q.3 In the Relational Model, the number of attributes and number of tuples in a relation are termed as _____ and _____ respectively.
(a) Cardinality, domain
(b) Degree, cardinality
(c) Domain, degree
(d) Cardinality, degree

Q.4 Making a change to the conceptual schema of a database but not affecting the existing external schemas is an example of
(a) Physical data independence
(b) Logical data independence
(c) Concurrency control
(d) Integrity control

Q.5 The data base administrator is, in effect, the coordinator between the _____ and the _____.
(a) DBMS, data base
(b) application program, data base
(c) data base, users
(d) application programs, users

Q.6 Consider the following ERD diagram illustrating the relationship of customers and banks.



Select from among the following, candidates for relations, if the above ERD is mapped into a relational model.
1. Customer(NID, CName)
2. Account(DateAcctCreated, BName, CName)
3. Bank(BankID, NID, BName)
4. Bank(BankID, BName)
5. Account(BankID, NID, DateAccCreated)

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

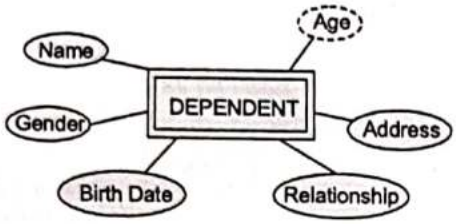
Q.7 Choose the incorrect statements.
1. In network model, data is represented by a collection of records, and relationship among data are represented by links.
2. In hierarchical model, data and relationships among data are represented by records and links respectively.
3. In hierarchical model, the records are organized as a collection of arbitrary graphs.
4. In network model, the records are organized as a collection of trees.

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

Q.8 Manager's salary details are hidden from the employee. This is

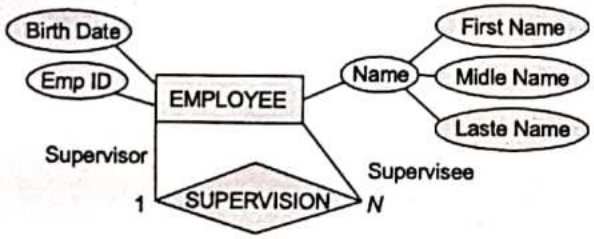
(a) conceptual level data hiding
(b) physical level data hiding
(c) external level data hiding
(d) none of the above

Q.9 The following diagram represents the dependent entity from an Entity Relationship diagram.



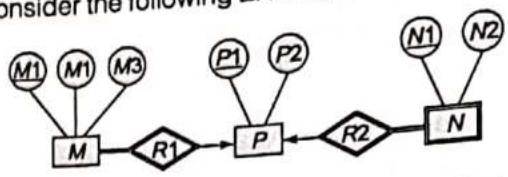
Select the characteristics which are not represented by the above diagram.
(a) BirthDate is a derived attribute
(b) Gender is an atomic attribute
(c) Address is a multivalued attribute
(d) Name is a key attribute

Q.10 Consider the following ERD diagram depicting the relationship of an employee and supervisor.



Which is 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.11 Consider the following ER diagram:



The minimum number of table needed to represent *M*, *N*, *P*, *R*1, *R*2 is
(a) 2 (b) 3
(c) 4 (d) 5

Q.12 Match List-I with List-II and select the correct answer using the codes given below the lists:

List-I		List-II
A.		1. Identifying relationship
B.		2. Weak entity
C.		3. Derived attribute
D.		4. Multivalued attribute

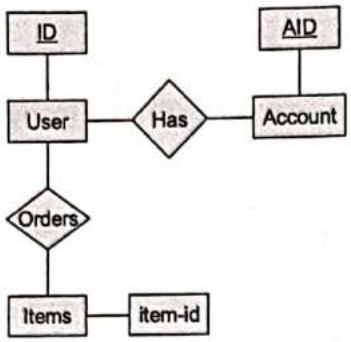
Codes:

	A	B	C	D
(a)	1	3	4	2
(b)	2	4	3	1
(c)	2	3	4	1
(d)	1	4	3	2

Q.13 Which of the following is true with respect to generalization:
1. It is the process of identifying subsets of an entity set that share some common characteristics but still contains some differences.
2. It tells that a class inherits properties of base class.

(a) Only 1
(b) Only 2
(c) Both 1 and 2
(d) None of these

Q.14 Consider the following ER model:



How many minimum tables are required for this *ER* model?

Q.15 R is relationship with 1 : 1 cardinality, 30% participation at E_1 end 70% participation at E_2 end which is the best possible design?

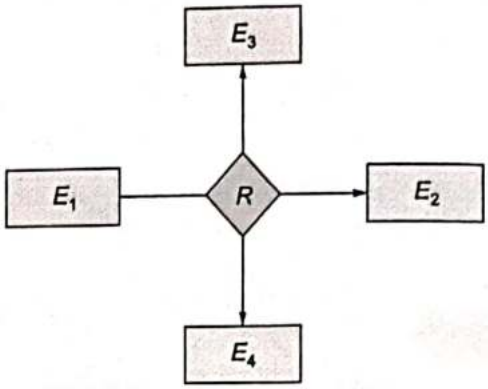
(a) E_1 and E_2 kept separate with foreign key at E_1 end

(b) E_1 and E_2 kept separate with foreign key at E_2 end

(c) E_1 and E_2 kept separate with foreign key at E_1 as well as E_2

(d) E_1 and E_2 merges into a single table with no foreign key

Q.16 Consider the following *ER* diagram with three entity sets E_1, E_2, E_3 and relationship set R



If E_1, E_2 and E_3 has 50, 30, 100 and 400 records respectively. What is the maximum number of records of entities that could be in the relationship set R?

■■■■

Answers The Relational Model

1.	(c)	2.	(b)	3.	(b)	4.	(b)	5.	(c)	6.	(b)	7.	(d)	8.	(c)	9.	(d)
10.	(d)	11.	(a)	12.	(c)	13.	(c)	15.	(b)								

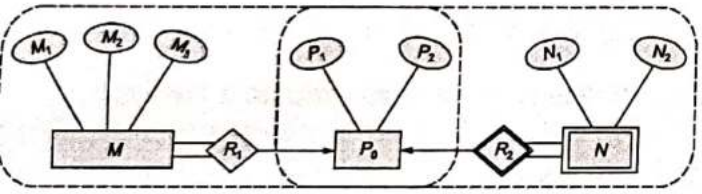
Explanations The Relational Model

1. (c)
Primary key column never contains NULL.
∴ Option (c) can not be used.
Remaining options can use NULL.

9. (d)
Ellipse with underlined attribute is known as key attribute.

10. (d)
The given ERD shows the self recursively relationship among Employees.
1 is supervisor 2 is supervisee.
The possible relation is
Employee (Emp Id, Birthdate, Salary, Name (First Name, Middle Name, Last Name), Supervisor Id)

11. (a)
Minimum number of tables needed is 2 as



1st table have the relation between *M* and *P* and 2nd table as *P* and *N*.
Hence option (a) is correct.

14. (3)
1. User (ID)
2. Account (ID, A.ID)
3. Items (Item_id, ID)

15. (b)
If E_1 is merged with R it contains 70% NULL values.
If E_2 is merged with R it contains 30% NULL values.
Merging which gives less number of Null value is preffered.
E_1 and E_2 can't be merged into a single table because E_1RE_2 has NO primary key

16. (50)
For relationship set R candidate and E_1 candidate key is same because between E_1 to E_3, E_1 to E_2 and E_1 to E_4 there is one to many relationship.

■■■■

2 CHAPTER
Database Design and Normalization

Q.1 Given FD = $\{AB \to C, CD \to E, E \to FG, G \to H\}$. Find which of the following can be a possible deduction from given functional dependencies.
(a) $DFG \to E$ (b) $ACD \to B$
(c) $ABD \to E$ (d) None of these

Q.2 Consider a relation 'R', all of its candidate keys are singletons. If R is in 3NF then R is _____.
(a) BCNF
(b) Not BCNF
(c) May not be BCNF
(d) Information is not sufficient

Q.3 Consider R = ABCDEFGH and the following FD's:
$H \to GD$
$E \to D$
$HD \to CE$
$BD \to A$
Identify the minimal cover of the given FD's?
(a) $\{H \to G, E \to D, H \to C, HD \to E, BD \to A\}$
(b) $\{H \to G, H \to D, E \to D, H \to C, BD \to A\}$
(c) $\{H \to G, E \to D, H \to C, HD \to C, BD \to A\}$
(d) $\{H \to G, E \to D, H \to C, H \to E, BD \to A\}$

Q.4 Consider a relation $R(A\ B\ C\ D\ E)$ with sets of FD's $F\{A \to B, BC \to E, ED \to A\}$ Given R is in which highest normal forms?
(a) 1 NF (b) 2 NF
(c) 3 NF (d) BCNF

Q.5 Let R be the relational schema with FDs F, is decomposed into R_1 and R_2 is loss less join decomposition only if
(a) (i) $R_1 \cup R_2 \subset R$
(ii) $R_1 \cap R_2 \neq \phi$
(iii) $R_1 \cap R_2$ gives candidate key that should be primary key in either R_1 or R_2
(b) (i) $R_1 \cup R_2 \equiv R$
(ii) $R_1 \cap R_2 \neq \phi$
(iii) $R_1 \cap R_2$ gives candidate key that should be primary key in either R_1 or R_2
(c) (i) $R_1 \cup R_2 \equiv R$
(ii) $R_1 \cap R_2 \neq \phi$
(iii) $R_1 \cap R_2$ gives candidate key that should not be primary key in either R_1 or R_2
(d) (i) $R_1 \cup R_2 \not\equiv R$
(ii) $R_1 \cap R_2 \neq \phi$
(iii) $R_1 \cap R_2$ gives candidate key that should not be primary key in either R_1 or R_2

Q.6 Consider the relation schema $S = \{A, B, C, D\}$ and the following functional dependencies on S.
$A \to BCD$
$B \to C$
$CD \to A$
Which of the following is true?
(a) S is in 3 NF and also in BCNF
(b) S is in 3 NF but not in BCNF
(c) S is in 2 NF but not in 3 NF
(d) S is in BCNF but not in 3 NF

Q.7 A relation R with two attribute is always in
(a) 1 NF (b) 2 NF
(c) 3 NF (d) BCNF

Q.8 Consider the universal relation $R = \{A, B, C, D, E, F, G, H, I, J\}$ and the set of FD, $F = \{AB \to C, A \to DE, B \to F, F \to GH, D \to IJ\}$. R is decomposed into
$R_1 = \{A, B, C, D, E\}$
$R_2 = \{B, F, G, H\}$
$R_3 = \{D, I, J\}$
Then the decomposition of R is
(a) Lossless join and dependency preservation
(b) Lossy but dependency preservation
(c) Lossless join but not dependency preservation
(d) Lossy and loss of dependency

Q.9 Consider the following relation:
Vehicle (vehicle No, sold_date, salesmanNo, commission%, discount_amt)
Assume that a vehicle may be sold by multiple salesman and vehicle {vehicle No, salesmanNo} is Primary key. Additional dependencies are Sold_date → discount_amt and salesmanNo → commission%.
Given relation is in which normal form?
(a) 1 NF (b) 2 NF
(c) 3 NF (d) None of these

Q.10 Given the relational schema $R(A, B, C, D, E)$ with the following functional dependencies.
$A \rightarrow B, BD \rightarrow E, C \rightarrow AD, E \rightarrow C$
Find which of the following is not a candidate key?
(a) E (b) C
(c) BD (d) A

Q.11 Find the number of FD's in the minimal cover of following FD set $FD = \{A \rightarrow B, CD \rightarrow A, CB \rightarrow D, CE \rightarrow D, AE \rightarrow F, AC \rightarrow D\}$.
(a) 2 (b) 3
(c) 4 (d) 5

Q.12 Find the number of possibilities to choose a primary key of the relation $R(A, B, C, D, E)$.
(a) 5 (b) 25
(c) 31 (d) 120

Q.13 Let R is a relation schema, $R(A, B, C, D)$ and $F = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$ is the set of functional dependency. How many candidate keys are there?
(a) 1 (b) 2
(c) 3 (d) None of these

Q.14 Suppose we have a relation $R(A, B, C)$ with multivalued dependency $\{A \rightarrow\rightarrow B\}$. If we know that tuples $(a, b1, c1)$, $(a, b2, c2,)$, $(a, b2, c3,)$, $(a, b3, c3,)$ are in R what other tuples do we know must also be in R
(a) $(a, b1, c2)$, $(a, b1, c3)$, $(a, b2, c1)$
(b) $(a, b1, c2)$, $(a, b2, c1)$, $(a, b3, c1)$, $(a, b3, c1)$, $(a, b3, c2)$
(c) $(a, b1, c2)$, $(a, b1, c3)$, $(a, b2, c1)$, $(a, b3, c1)$, $(a, b3, c2)$
(d) None of the above

Q.15 Consider the relation given below and find the maximum normal form applicable to them
1. $R(A, B)$ with productions $\{A \rightarrow B\}$
2. $R(A, B)$ with productions $\{B \rightarrow A\}$
3. $R(A, B)$ with productions $\{A \rightarrow B, B \rightarrow A\}$
4. $R(A, B, C)$ with productions $\{A \rightarrow B, B \rightarrow A, AB \rightarrow C\}$

(a) 1, 2 and 3 are in 3NF and 4 is in BCNF
(b) 1 and 2 are in BCNF and 3 and 4 are in 3NF
(c) All are in 3NF
(d) All are in BCNF

Q.16 Consider the relation $R(A, B, C, D, E)$ and the set
$F = \{AB \rightarrow CE, E \rightarrow AB, C \rightarrow D\}$
What is the highest normal form of this relation?
(a) 1NF (b) 2NF
(c) 3NF (d) BCNF

Q.17 Given the relation $R(X, Y, W, Z, P, Q)$ and the set
$F = \{XY \rightarrow W, XW \rightarrow P, PQ \rightarrow Z, XY \rightarrow Q\}$
Consider the decomposition $R_1(Z, P, Q)$ and $R_2(X, Y, W, P, Q)$. This decomposition is
(a) Lossless decomposition
(b) Lossy decomposition
(c) Either lossless or lossy
(d) Neither lossless or lossy

Q.18 Consider the relation SCHEDULE shown below. What is the highest normal form of this relation?
SCHEDULE (Student_ID, Class_No, Student_Name, Student_Major, Class_Time, Building _Room, Instructor).
Assume the following functional dependencies {Student_ID → Student_Name, Student_ID → Student_Major, Class_No → Class_time, Class_No → Building Room, Class_No→ Instructor}
(a) 1NF (b) 2NF
(c) 3NF (d) BCNF

Q.19 A relation R is defined as $R(S\#, STATUS, CITY, SNAME)$ where $S\#$ is the primary key. If R is decomposed into two relations R_1 and R_2, which of the following is a loss less decomposition?
(a) $R_1(S\#, STATUS)$, $R_2(S\#, CITY, SNAME)$
(b) $R_1(S\#, STATUS)$, $R_2(STATUS, CITY, SNAME)$

(c) R_1(S#, STATUS, CITY), R_2 (CITY, SNAME)
(d) R_1(S#, STATUS, SNAME), R_2(CITY, STATUS)

Q.20 Consider the relation $R(S, T, U, V)$ and the set $F = \{S \to T, T \to U, U \to V$ and $V \to S\}$. Let R is decomposed into R_1 and R_2 such that $R_1 \cap R_2 = \phi$, the decomposition is
(a) Lossless decomposition
(b) Lossy decomposition
(c) Dependency preservation
(d) Cannot say

Q.21 Suppose that we decompose the schema $R(A, B, C, D, E)$ into $R_1(A, B, C)$ and $R_2(A, D, E)$ if the set F are functional dependencies holds on R.

$F = \{A \to BC, CD \to E, B \to D, E \to A\}$

The decomposition is a
(a) Lossless decomposition
(b) Lossy decomposition
(c) May or may not be lossless decomposition
(d) None of these

Q.22 Find the highest normal form of the relation $R(A, B, C, D)$ and the set

$F = \{AB \to D, AC \to BD, B \to C\}$

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

Q.23 Consider the following set of functional dependencies (*FD*s)
{Emp_No → Ename, Bdate, Address, Dept_No}
{Dept_No → Dname, Mgr_No}

The additional *FD*(s) which cannot be inferred from the above set of *FD*s is
(a) {Emp_No → Dname, Mgr_No}
(b) {Ename → Dept_No}
(c) {Emp_No → Dname}
(d) {Ename, Dept_No → Mgr_No}

Q.24 A relation (from the relational database model) consists of a set of tuples, which implies that
(a) all tuples in a relation must be distinct
(b) relational model supports multi-valued attributes whose values can be represented in sets
(c) for any two tuples, the values associated with all of their attributes may be the same
(d) all tuples in a particular relation may have different attributes

Q.25 If K is a foreign key in relation $R1$, then
(a) Every tuple of R1 has a distinct value for K
(b) K cannot have a null value for tuples in $R1$
(c) K is a key for some other relation
(d) K is a primary key for $R1$

Q.26 If W, X, Y and Z are attributes of a relation, which of the following inference rules for functional dependencies is correct?
(a) If $Y \subset X$ then $Y \to X$
(b) If $X \to Y$ then $(X, Z) \to (Y, W)$
(c) If $X \to (Y, Z)$ then $X \to Y$
(d) If $(X, Z) \to Y$ then $X \to Y$ and $Z \to Y$

Q.27 Consider relation schema $R(A, B, C, D, E)$ with functional dependencies F containing $A \to B$ and $C \to D$. The schema is decomposed into $R_1(A, B)$ and $R_2(A, C, D, E)$. Find which relation is not in BCNF.
(a) R_1 is not in BCNF
(b) R_2 is not in BCNF
(c) R_1 is not in BCNF but R_2 is in BCNF
(d) None of these

Q.28 Let R(ABCDEFGH) be a relation with the following functional dependencies

$FD = \{AB \to CDEF$
$A \to C$
$B \to D$
$C \to FG$
$G \to H$
$H \to F$
$G \to F\}$

Find the maximum normal form applicable to R.
(a) 1NF (b) 2NF
(c) 3NF (d) BCNF

Q.29 Consider the following functional dependencies in a database:

Date 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 3 NF
(b) in 3NF but not in BCNF
(c) in BCNF
(d) None of the above

Q.30 Consider the following instance:

X	Y	Z
1	4	2
1	5	3
1	6	3
3	2	2

Which one of the following correctly describe the functional dependency hold by above instance.
(a) $XY \rightarrow Z$ and $Z \rightarrow Y$
(b) $YZ \rightarrow X$, $Y \rightarrow Z$ and $Y \rightarrow X$
(c) $YZ \rightarrow X$ and $X \rightarrow Z$
(d) $XY \rightarrow Y$ and $X \rightarrow Z$

Q.31 Consider the following set of functional dependencies F on the schema (A, B, C)
$F = \{A \rightarrow BC, B \rightarrow C, A \rightarrow B, AB \rightarrow C\}$
The canonical cover of this set is
(a) $A \rightarrow BC$ and $B \rightarrow C$
(b) $A \rightarrow BC$ and $AB \rightarrow C$
(c) $A \rightarrow BC$ and $A \rightarrow B$
(d) $A \rightarrow B$ and $B \rightarrow C$

Q.32 $f = \{A \rightarrow BC, B \rightarrow A, C \rightarrow A\}$
$g = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$
(a) f covers g
(b) g covers f
(c) f and g are equivalent
(d) None of these

Q.33 Given a relation $R(A, B, C, D, E)$ with functional dependencies $A \rightarrow B$, $B \rightarrow C$ and $BD \rightarrow E$. Which one of the following could be result of decomposition of R into BCNF?
(a) AB, BC, BDE (b) AD, BDE, ABC
(c) ABD, BE, CE (d) AB, ACD, BCE

Q.34 Consider the relation $R(A, B, C, D, E)$ with dependencies $A \rightarrow\rightarrow BC$ and $D \rightarrow B$ of the following dependencies.
1. $A \rightarrow B$
2. $A \rightarrow\rightarrow C$
3. $A \rightarrow\rightarrow D$

Which must necessarily hold in R?
(a) 1 only (b) 2 only
(c) 2 and 3 (d) 1, 2 and 3

Q.35 Cnsider the relation schema $S(A, B, C, D)$ and the following functional dependencies on
$S\{A \rightarrow BCD, B \rightarrow C, CD \rightarrow A\}$
If we decompose S into $S_1 = ABD$ and $S_2 = BC$, then fill in the following blanks for A, B and C
1. The decomposition is _____.
2. The decomposition is _____.
3. Valid decomposition into _____.

(a) lossless, dependency preserving, BCNF
(b) lossless, not dependency preserving, BCNF
(c) lossy, not dependency preserving, 3NF
(d) lossy, dependency preserving, BCNF

Q.36 Consider the following table:

A	B	C
x	y	z
m	n	n
n	m	z
o	p	p
q	q	q
x	z	n

Which of the following dependencies holds true for the above table?
(a) $AB \rightarrow C$ (b) $BC \rightarrow A$
(c) Both (a) and (b) (d) None of these

Q.37 In the relation $R(A, B, C, D)$ with functional dependencies $F = \{AB \rightarrow CD, C \rightarrow A, D \rightarrow B\}$ How many number of candidate keys are there for the above relation?
(a) 2 (b) 3
(c) 4 (d) None of these

Q.38 Consider the following relation and its sample data. (Consider that these are the only tuples for the given relation)?

Emp No	Dept No	Proj No
1001	01	12
1001	01	13
1002	01	12
1003	01	14

1. The functional dependency (Emp No, Dept No)→ ProjNo holds over *R*.
2. The functional dependency DeptNo → ProjNo holds over *R*.
3. The functional dependency EmpNo → DeptNo holds over *R*.
4. The functional dependency Proj No → DeptNo holds over *R*.
5. The functional dependency (EmpNo, ProjNo) → DeptNo holds over *R*.

(a) 1, 2 and 3 (b) 2, 3 and 4
(c) 3, 4 and 5 (d) All of these

Q.39 Consider the relation 'Property (PropertyId Village name, Lot# Area)' with the following functional dependencies.

1. PropertyId → (VillageName, Lot#, Area)
2. (VillageName, Lot#) → (PropertyId, Area)
3. Area → VillageName

Which one of the following statements is false with respect to the information given above?

(a) The relation 'Property' is in 3NF.
(b) The functional dependency Area → Village Name violates BCNF.
(c) The relation 'Property' is in BCNF.
(d) All of the above.

Q.40 The terms in List-I have been mapped to List-II so that it corresponds to the mapping process of the ER Model into a Relational model. Which of the following represents the mapping process?

List-I

A. Entity type
B. Key attributes
C. Composite attribute
D. Multivalued attribute
E. Value set

List-II

1. Primary or (Secondary) Key
2. Domain
3. Relation and Foreign key
4. Set of simple component attributes
5. Relation

Codes:

	A	B	C	D	E
(a)	3	1	4	2	5
(b)	5	1	4	3	2
(c)	3	1	4	5	2
(d)	5	1	3	4	2

Q.41 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 are correct.

(a) Only 1 and 3 (b) Only 2 and 3
(c) Only 1 and 2 (d) All of these

Q.42 Which of the following statements is not correct with respect to normalization?

(a) Normalization is a formal technique that can be used at any stage of the database design.
(b) Normalization can be used as a validation technique to check the structure of relations which may have been created using ER modeling.
(c) The process of normalization through decomposition must achieve the lossless join property at any cost whereas the dependency preservation property is sometimes sacrificed.
(d) None of the above.

Q.43 Choose the correct statements.

(a) An alternate key is a candidate key that is not a primary key
(b) An alternate key is a primary key that is not a candidate key
(c) An alternate key is a candidate key that is also a primary key
(d) None of the above

Q.44 Third normal form is inadequate in situations where the relation

(a) has multiple candidate keys
(b) has candidate keys that are composite
(c) has overlapped candidate keys
(d) All of the above

Q.45 Redundancy is dangerous as it is a potential threat to data
(a) integrity (b) consistency
(c) sufficiency (d) both (a) and (b)

Q.46 For relation $R(A, B, C, D, E, F)$ the set of *FDs* is $\{A \rightarrow C, B \rightarrow D, C \rightarrow E, D \rightarrow E, E \rightarrow A, F \rightarrow B\}$
What is the candidate key for R
(a) A (b) C
(c) D (d) F

Q.47 Let $R(a, b, c)$ and $S(d, e, f)$ be two relations in which d is the foreign key of S that refers to the primary key of R. Consider the following four operations R and S
1. Insert into R
2. Insert into S
3. Delete from R
4. Delete from S

Which of the following is true about the referential integrity constraint above?
(a) None of 1, 2, 3 or 4 can cause its violation
(b) All of 1, 2, 3 and 4 can cause its violation
(c) Both 1 and 4 can cause its violation
(d) Both 2 and 3 can cause its violation

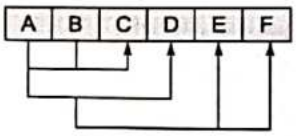
Q.48 A set of *FDs* is given for a relation $R(A, B, C, D, E, F, G)$
$\{A \rightarrow B, BC \rightarrow DE, AEF \rightarrow G\}$
What is the closure of $\{A, C\}^+$ under this set?
(a) A, B, C (b) A, B, C, D
(c) A, B, C, D, E, F (d) A, B, C, D, E

Q.49 Which the relation R with four attributes A, B, C and D the functional dependencies
$\{(A, B) \rightarrow (C, D)$ and
$C \rightarrow D\}$
Which of the following statements is/are correct?
(a) C is key for relation R
(b) R is in 3 NF
(c) Functional dependency $C \rightarrow D$ violates 3 NF
(d) R is in BCNF

Q.50 Consider two relation schema $R_1 = (A, B, C, D, E)$ and $R_2 (A, B, C, D, E)$. Statement 1 is the function dependency of R_1 and 2 is the functional dependency of R_2.
1. $A \rightarrow B, AB \rightarrow C, D \rightarrow AC, D \rightarrow E$
2. $A \rightarrow BC, D \rightarrow AE$

Which of the following statement is true?
(a) Functional dependency of R_1 is equivalent to functional dependency of R_2.
(b) Functional dependency of R_1 and R_2 are not equivalent.
(c) We can't compare functional dependency of R_1 and R_2.
(d) None of the above

Q.51 For the given relation instance the best normal form is



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

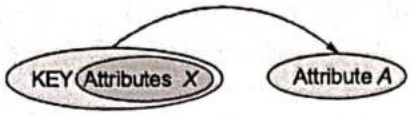
Q.52 Consider a schema $R(A, B, C, D)$ and functional dependencies: $A \rightarrow B$ and $C \rightarrow D$. Then the decomposition of R into $R_1 (AB)$ and $R_2 (CD)$ is
(a) dependency preserving and lossless join
(b) lossless join but not dependency preserving
(c) dependency preserving but not lossless join
(d) not dependency preserving and not lossless join

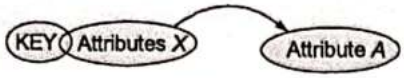
Q.53 Let R be a relation. Which of the following comments about the relation R are correct?
1. R will necessarily have a composite key if R is in BCNF but not in 4NF
2. If R is in 3NF and if every key of R is simple, then R is in BCNF
3. If R is in BCNF and if R has at least one simple key, then R is in 4NF
4. If R is in 3NF and if its every key is simple, then R is in 5NF

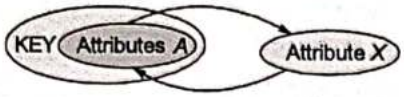
(a) 1 and 4 (b) 2 and 3
(c) 1, 2 and 3 (d) 1, 2, 3 and 4

Q.54 Match List-I with List-II and select the correct answer using the codes given below the lists:

List-I

A. 

B. 

C. 

List-II

1. Transitive dependencies.
2. Trivial dependencies.
3. Partial dependencies.

Codes:

	A	B	C
(a)	1	2	3
(b)	1	3	1
(c)	3	1	2
(d)	3	2	1

Q.55 From the following instance of a relation schema $R(A, B, C)$, we can conclude that:

A	B	C
1	1	1
1	1	0
2	3	2
2	3	2

(a) A functionally determines B and B functionally determines C
(b) A functionally determines B and B does not functionally determine C
(c) A does not functionally determine C
(d) All of the above are correct

■■■■

Answers Database Design and Normalization

1.	(c)	2.	(a)	3.	(d)	4.	(c)	5.	(b)	6.	(b)	7.	(d)	8.	(a)	9.	(d)
10.	(a)	11.	(d)	12.	(c)	13.	(c)	14.	(c)	15.	(d)	16.	(b)	17.	(a)	18.	(a)
19.	(a)	20.	(b)	21.	(a)	22.	(c)	23.	(b)	24.	(a)	25.	(c)	26.	(c)	27.	(b)
28.	(a)	29.	(d)	30.	(b)	31.	(d)	32.	(c)	33.	(a)	34.	(b)	35.	(b)	36.	(c)
37.	(c)	38.	(c)	39.	(c)	40.	(b)	41.	(c)	42.	(d)	43.	(a)	44.	(d)	45.	(d)
46.	(d)	47.	(d)	48.	(d)	49.	(c)	50.	(a)	51.	(a)	52.	(c)	53.	(d)	54.	(c)
55.	(d)																

Explanations Database Design and Normalization

1. (c)

$AB \rightarrow C$

Augment 'D'

$ABD \rightarrow CD$

and we have $CD \rightarrow E$.

$\therefore$ Transitively : $ABD \rightarrow E$ is a possible deduction.

OR

$(ABD)^+ = ABCDEFGH$

$\therefore$ $ABD \rightarrow E$ is possible

2. (a)

Given R is in 3NF.

Either FD LHS is key or RHS is prime attributes.

If LHS FD is a key then it is BCNF

If RHS FD is prime attribute then it must be key, because all candidate keys are singleton. So it is also in BCNF.

$\therefore$ R is also in BCNF

3. **(d)**

Given *FD*s:

$H \rightarrow GD$
$E \rightarrow D$
$HD \rightarrow CE$
$BD \rightarrow A$

(1) Right reduced *FD*s:

$H \rightarrow G$
$H \rightarrow D$
$E \rightarrow D$
$HD \rightarrow C$
$HD \rightarrow E$
$BD \rightarrow A$

(2) Left reduced *FD*s:

$H \rightarrow G$
$H \rightarrow D$
$E \rightarrow D$
$H \rightarrow C$
$H \rightarrow E$
$BD \rightarrow A$

(3) Minimal cover:

$H \rightarrow G$
$E \rightarrow D$
$H \rightarrow C$
$H \rightarrow E$
$BD \rightarrow A$

4. **(c)**

$F = \{A \rightarrow B, BC \rightarrow E, ED \rightarrow A\}$
$A^+ = A, B,$
$(BC)^+ = B, C, E$
$(ED)^+ = E, D, A, B$

So candidate keys may be {*EDC*, *BCD*, *ACD*} and prime attributes are {*A*, *B*, *C*, *D*, *E*}

BCNF: If $x \rightarrow y$ then x should be candidate key then the relation will be in BCNF. In the given *FD*s set $\{A \rightarrow B, BC \rightarrow E, ED \rightarrow A\}$, x may be *A* or *BC* or *ED* which are not candidate key so given relation is not in BCNF.

3NF: If $x \rightarrow y$ then (i) x should be candidate key, or (ii) y should be prime attribute in our given relation y may be *B*, *E*, *A*, which are prime attributes. So given relation is in 3 NF.

6. **(b)**

$F = \{A \rightarrow BCD, B \rightarrow C, CD \rightarrow A\}$
$A^+ = A\ B\ C\ D$
$CD^+ = A\ B\ C\ D$

So candidate keys are {*A*, *CD*}
Prime attributes are {*A*, *C*, *D*}

Check for BCNF :

$A \rightarrow BCD$ ✓ ($\because$ *A* is candidate key)
$B \rightarrow C$ X ($\because$ *B* is not candidate key)
$CD \rightarrow A$ ✓ (*CD* is candidate key)

Check for 3NF :

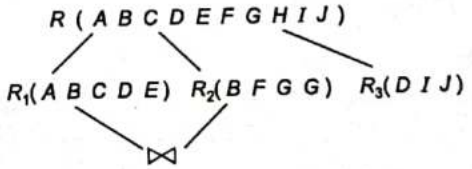
$A \rightarrow BCD$ ✓ ($\because$ *A* is candidate key)
$B \rightarrow C$ ✓ ($\because$ *C* is prime attribute)
$CD \rightarrow A$ ✓ ($\because$ *CD* is candidate key)

7. **(d)**

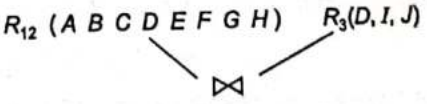
Let the relation is *R*(*A*, *B*)
if *A* is key then $A \rightarrow B$ which is in BCNF
if *B* is key then $B \rightarrow A$ which is also in BCNF
if *AB* is key then $AB \rightarrow AB$ which is in BCNF
so *R*(*A*, *B*) will always in BCNF

8. **(a)**

$F = \{AB \rightarrow C, A \rightarrow DE, B \rightarrow F, F \rightarrow GH, D \rightarrow IJ\}$
$(AB)^+ = \{A\ B\ C\ D\ E\ F\ G\ H\ I\ J\}$
so AB is candidate key



Common attribute between R_1 and R_2 is *B*
$\because B^+ = B, F, G, H$
so *B* is c.key in R_2



Common attribute between R_{12} and R_3 is *D*
D, and $D^+ = D, I, J$
so *D* is c. key in R_3

So $R_{123}(A\ B\ C\ D\ E\ F\ G\ H\ I\ J)$
The this decomposition is loss less decomposition.
Check for dependency preservation

$R(A\,B\,C\,D\,E\,F\,G\,H\,I\,J)$

$R_1(A, B, C, D, E)$ $F_1 = \{AB \to CDE\}$ $\{A \to DE\}$	$A^+ = A, D, E$ $B^+ = B$ $(AB)^+ = A, B, C, D, E$
$R_2(B, F, G, H)$ $F_2 = \{B \to FGH\}$ $\{F \to GH\}$	$B^+ = B, F, G, H$ $F^+ = F, G, H$
$R_3(D, I, G)$ $F_2 = \{D \to IJ\}$	$D^+ = D, I, G$

$$F_1 \cup F_2 \cup F_3 = \begin{Bmatrix} AB \to CDE, & B \to FGH, & D \to IJ \\ A \to DE, & F \to GH & \end{Bmatrix}$$

$$\equiv F$$

So $F_1 \cup F_2 \cup F_3$ is equivalent to F.
so dependency is preserved.

9. (d)
In the given relation salesman No. has multiple attribute so it is not in 1NF.

10. (a)

$$(E)^+ = ECADB = ABCDE$$
$$(C)^+ = ABCDE$$
$$(BD)^+ = ABCDE$$

$(A)^+$ can not determine all attributes.
$\therefore$ A is not a candidate key.

11. (d)
$AC \to D$ can be eliminated, it can be derived from $A \to B$ and $CB \to D$ using augmentation and transitive rule.

$$A \to B \Rightarrow AC \to BC$$
$$\Rightarrow AC \to D$$

and remaining FDs are not possible to eliminate.
$\therefore$ 5 FDs are there in the minimal cover.

12. (c)
Number of possible primary keys are :
$2^5 - 1 = 31$ (with zero attributes not possible)

13. (c)
The candidate keys are AD, BD and CD.

14. (c)
If the $MD\ A \to\to B$ holds, then the $MD\ A \to\to C$ also holds. Notice that all the A values are the same.

We are given

a	b_1	C_1
a	b_2	C_2
a	b_2	C_3
a	b_3	C_3

Since there are three B values, and three C values (further same A), we can confidently state that there will be nine combinations. The question only lists four, we can thus state the following five tuples.

a	b_1	C_2
a	b_1	C_3
a	b_2	C_1
a	b_3	C_1
a	b_3	C_2

15. (d)
Consider the relations 1, 2 and 3. All are binary relations. A binary relation always be in BCNF. Coming to 4th relation it is also in BCNF because A, B and AB all are super keys.

16. (b)
Two possible keys of the relation are AB and E. Therefore prime attributes are, A, B and E. Non prime attributes are C and D. The relation is in 2NF because there are no partial dependency on any of the key.
The relation is not in 3NF because of $C \to D$ functional dependency.

17. (a)
$R_1 \cap R_2 \to R_1$
i.e. $(ZPQ) \cap (XYWPQ) = PQ$ and PQ is Primary key of R_1.

18. (a)
There are some partial dependences on the key, for example, Class_No → Building_Room or Student_ID → Student_Major. Therefore the highest normal form of the above relation is 1NF.

19. (a)
$R_1 \cap R_2 = S\#$ which is primary key in R_1.

21. (a)
$R_1 \cap R_2 \to R_1$ i.e. $(ABC) \cap (ADE) = A$ and A is Primary key in R_1.

22. (c)

The key of this relation are AB and AC. The prime attributes are $\{A, B, C\}$. So, the relation is in 3NF.

23. (b)

$\{Emp\_No\}^+ = \{Ename, Bdate, Address, Dept\_No, Dname, Mgr\_No\}$

$\{Ename, Dept\_No\}^+ = \{Ename, Dname, Mgr\_No\}$

$\{Ename\}^+ = \{Ename\}$

27. (b)

$R(A, B, C, D)$ is decomposed into

$R_1(A, B)$ with $FD = \{A \rightarrow B\}$ and

$R_2(A, C, D, E)$ with $FD = \{C \rightarrow D\}$

For R_1: A is superkey so R_1 is in BCNF

For R_2 : C is not a superkey.

Hence R_2 is not in BCNF.

28. (a)

The primary key of relation is AB. And there exist partial dependencies $A \rightarrow C$, $B \rightarrow D$. Hence Relation is not in 2NF.

29. (d)

For this relation, only following FDs are of consideration

1. Date of Birth → Age
2. Name → Roll_number
3. Roll_number → Name

Keys of relation are (name, Date of Birth) and (Roll_number, Date of Birth). The relation is not in 2NF because of partial dependency i.e. Date of Birth → Age. The other two FDs ae not partial bacause their right hand side attributes are prime attributes.

30. (b)

$YZ \rightarrow X$ is true because YZ is not repeating anywhere. If YZ repeats then X has to be repeated.

$Z \rightarrow Y$ is false because in first tuple when value of Z is 2, then value of Y is 4, but in last tuple value of Y has to be 4 but it is 2. Similarly is the care with 2nd and 3rd tuple.

31. (d)

$$\text{Old set} \begin{cases} A \rightarrow BC \\ B \rightarrow C \\ A \rightarrow B \\ AB \rightarrow C \end{cases}$$

Now check extraneous in L.H.S. Take B as extraneous attribute

$$\text{New set} \begin{cases} A \rightarrow BC \\ B \rightarrow C \\ A \rightarrow B \\ A \rightarrow C \end{cases}$$

Now old FD's imply the new FDs

So B is an extraneous attribute.

Similarly if we check for C, it is also an extraneous attribute. But we will remove only one attribute to preserve the dependency in the set.

Now the reduced FD's will be

$$\left.\begin{array}{l} A \rightarrow BC \\ B \rightarrow C \\ A \rightarrow B \\ A \rightarrow C \end{array}\right\}$$

But the last two dependencies can be removed because they can be derived from $A \rightarrow BC$. Now, the FD set will be

$$\left.\begin{array}{l} A \rightarrow BC \\ B \rightarrow C \end{array}\right\} \text{old}$$

now we check extraneous attribute in R.H.S. Take C as extraneous

$$\left.\begin{array}{l} A \rightarrow B \\ B \rightarrow C \end{array}\right\} \text{new}$$

here new set of FD's implies old set of FD's. So C is an extraneous attribute. Similarly B is not an extraneous attribute.

32. (c)

$f = \{A \rightarrow BC, B \rightarrow A, C \rightarrow A\}$

$g = \{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$

To check if f covers g, every FD in g logically implies in f.

FD in g : $A \rightarrow B$, $B \rightarrow C$, $C \rightarrow A$

Check for $A \rightarrow B$,

$A^+ \rightarrow ABC$ $\{A \rightarrow BC$ in $f\}$

Check for $B \rightarrow C$,

$B^+ \rightarrow BAC$, $\{B \rightarrow A, A \rightarrow BC$ in $f\}$
Check for $C \rightarrow A$,
$C^+ \rightarrow CAB$ $\{C \rightarrow A, A \rightarrow BC$ in $f\}$
Therefore, f covers g.
To check if g covers f:
FD in f: $A \rightarrow BC$, $B \rightarrow A$, $C \rightarrow A$
Check for $A \rightarrow BC$,
$A^+ \rightarrow ABC$ $\{A \rightarrow B, B \rightarrow C$ in $g\}$
Check for $B \rightarrow A$,
$B^+ \rightarrow BCA$ $\{B \rightarrow C, C \rightarrow A$ in $g\}$
Check for $C \rightarrow A$,
$C^+ \rightarrow A$ $\{C \rightarrow A$ in $g\}$
Therefore, g covers f.
Both f and g covers each other, so they are equivalent $f \equiv g$.

33. (a)
All the FDs violate BCNF, so create a separate relation for each functional dependency.

35. (b)
1. $S_1 \cap S_2 = B$ here $B \rightarrow C$. So lostless.
2. $CD \rightarrow A$ and $A \rightarrow C$ are lost.
3. Since FD $B \rightarrow C$ cause problem for BCNF.

So it is BCNF decomposition.

36. (c)
For each combination of attributes on LHS, there should be a unique combination on RHS.
It states true for both (a) and (b).

37. (c)
There are four keys AB, CD, AD, BC
$\{AB\}^+ = \{A\ B\ C\ D\}$
$\{CD\}^+ = \{A\ B\ C\ D\}$
$\{AD\}^+ = \{A\ B\ C\ D\}$
$\{BC\}^+ = \{A\ B\ C\ D\}$

38. (c)
The FD (EmpNo, DeptNo.) $\rightarrow$ ProjNo does not hold because for EmpNo 1001 and DeptNo of there are two ProjNo (12,13), which violates the FD.
Similarly FD DeptNo $\rightarrow$ ProjNo vilotes the FD.

39. (c)
The 'Property' Relation is not in BCNF bacause the FD Area $\rightarrow$ VillageName Vilotes BCNF.

47. (d)
Referential Integrity constraint
In relational model, two relation are related to each other over the basis of attributes. Every value of referencing attribute must be null or be available in the referenced attribute.

$R(\underline{a}, b, c)$ and $S(d, \underline{e}, f)$

here d is the foreign key of S that refers to the primary key of R.
1. Insert into R will not cause any voilation.
2. Insert into S may cause violation because for each entry in 'S' it must be in 'R'.
3. Delete from R may cause violation because for the deleted entry in R there may be referenced entry in the relation S.
4. Delete from S will not cause any violation.

Hence (d) is the correct option.

48. (d)
$R(A, B, C, D, E, F, G)$
As $A \rightarrow B$
$BC \rightarrow DE$
$\therefore (AC)^+ = \{A, B, C, D, E\}$
Hence (d) is the correct option.

49. (c)
$\{(A, B) \rightarrow (C, D)$ and $C \rightarrow D\}$
for $(A, B) \rightarrow (C, D)$
The candidate key is $A\ B$
also if $C \rightarrow D$ i.e. non key going for non-key hence C is neither a key and D is not a prime attribute hence $C \rightarrow D$ violating 3NF condition.
Hence (c) is correct option.

50. (a)
$R_1 = \{A, B, C, D, E\}$
$A \rightarrow B$, $AB \rightarrow C$, $D \rightarrow AC$, $D \rightarrow E$
$R_2 = \{A, B, C, D, E\}$
$A \rightarrow BC$, $D \rightarrow AE$
Here candidate key of relation R_1 is
$AD^+ = \{A, B, C, D, E\}$
FD'S set is $= \{A \rightarrow B, AB \rightarrow C, A \rightarrow C, D \rightarrow AC,$
$D \rightarrow A, D \rightarrow C, D \rightarrow E\}$
For relation R_2 candidate key is AD

FD's set is = $\{A \to B, A \to C, A \to BC, D \to A, D \to C, D \to E, D \to AC\}$
As we can see cleary $R \equiv R_2$
Hence option (a) is correct.

51. (a)
For the given relation instance we have the FD set as

$$AB \to C$$
$$A \to D$$
$$B \to EF$$

Hence candidate is AB
For

$AB \to C$ in BCNF
$A \to D$ in 1 NF (Partial dependency)
$B \to EF$ in 1 NF (Partial dependency) in BCNF
} 1NF

Hence (a) is the correct option

52. (c)
$R_1(AB)$, $R_2(CD)$ for $R(A, B, C, D)$
For losssless join, 3 conditions must hold by relations R_1 and R_3.

(1) $R_1 \cup R_2 = R$
$AB \cup CD = ABCD = R$ True

(2) $R_1 \cap R_2 \neq \phi$
$AB \cap CD = \phi$ False

So, not lossless join.
For dependency preserving,
In $R_1(AB)$ $A \to B$
$A^+ = AB$ A candidate key for R_1.
In $R_2(CD)$ $C \to D$
$C^+ = CD$ C candidate key for R_2.
So, the decomposition is dependency preserved but not lossless.

55. (d)
From the given instance we can determine that

when $A=1$ $B=1$
$A=2$ $B=3$

Hence A functionally determines B.

when $B=1$ $C=1$
$C=0$
$B=3$ $C=2$

Hence B does not functionally determining C.

■■■■

3 CHAPTER
Relational Algebra

Q.1 Which of the following relational algebraic operation is not a commutative operation?
(a) Union (b) Intersection
(c) Selection (d) Projection

Q.2 Consider a banking database with the following table with three attributes loans (br_name, loan_no, amount). Find the appropriate query for the given statements below.
"Find the loan number for each loan of an amount greater than 20000"
(a) $\{t \mid t \in \text{loans} \wedge t[\text{amount}] > 20000\}$ where 't' is a tuple
(b) $\{t \mid \exists\ s \in \text{loans}\ (t[\text{loan\_no}] = S[\text{loan\_no}] \wedge [\text{amount}] > 20000\}$
(c) $\{t \mid \forall\ s \in \text{loans}\ (t[\text{loan\_no}] = S[\text{loan\_no}] \wedge [\text{amount}] > 20000\}$
(d) None of the above

Q.3 Which of the following is wrong?
(a) $\pi_{L_1 \cup L_2}(E_1 \bowtie_\theta E_2) = (\pi_{L_1}(E_1)) \bowtie_\theta (\pi_{L_1}(E_2))$
(b) $\sigma_P(E_1 - E_2) = \sigma_P(E_1) - E_2$
(c) $\sigma_{\theta_1 \wedge \theta_2}(E) = \sigma_{\theta_1}(\sigma_{\theta_2}(E))$
(d) $E_1 \bowtie_\theta E_2 = E_2 \bowtie_{\theta_1} E_1$

Q.4 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?

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

(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

Q.5 Select the relational expression which could possibly return the following result.

a	c
1	2
2	3

(a) $\Pi_{a,c}(\sigma_{a=c}R)$ (b) $\Pi_{a<c}(\Pi_{a,c}R)$
(c) $\Pi_{a<2}R$ (d) $\sigma_{a<c}(\Pi_{a,c}R)$

Q.6 Consider a following declaration:
$r_1 = \pi_{\text{branch\_name}}(\sigma_{\text{branch\_city} = \text{"Dadar"}}(\text{branch}))$
$r_2 = \pi_{\text{cust\_name, branch\_name}}(\text{depositor} \bowtie \text{account})$
Then to find the customers who appears in r_2, with every branch name in r_1, the query is:
(a) $\pi_{\text{branch\_name}}(\sigma_{\text{branch\_city} = \text{"Dadar"}}(\text{branch})) \cap \pi_{\text{cust\_name, branch\_name}}(\text{depositor} \bowtie \text{account})$
(b) $\pi_{\text{cust\_name, branch\_name}}(\text{depositor} \bowtie \text{account}) \cap \pi_{\text{branch\_name}}(\sigma_{\text{branch\_city} = \text{"Dadar"}}(\text{branch}))$
(c) $\pi_{\text{branch\_name}}(\sigma_{\text{branch\_city} = \text{"Dadar"}}(\text{branch})) \div \pi_{\text{cust\_name, branch\_name}}(\text{depositor} \bowtie \text{account})$
(d) $\pi_{\text{cust\_name, branch\_name}}(\text{depositor} \bowtie \text{account}) \div \pi_{\text{branch\_name}}(\sigma_{\text{branch\_city} = \text{"Dadar"}}(\text{branch}))$

Q.7 Consider the following relation schemas:
b-Schema = (b-name, b-city, assets)
a-Schema = (a-num, b-name, bal)
d-Schema = (c-name, a-number)
Let branch, account depositor be respectively instance of the above schemas. Assume that account and depositor relations are much bigger than the branch relation.
Consider the following query:
$\Pi_{\text{c-name}}(\sigma_{\text{b-city} = \text{"Agra"} \wedge \text{bal} < 0}(\text{branch} \bowtie (\text{account} \bowtie \text{depositor})))$

Which one of the following queries is the most efficient version of the above query?

(a) $\Pi_{c\text{-name}}(\sigma_{bal < 0}(\sigma_{b\text{-city} = \text{"Agra"}}\ branch \bowtie account \bowtie depositor)$

(b) $\Pi_{c\text{-name}}(\sigma_{b\text{-city} = \text{"Agra"}}\ branch \bowtie (\sigma_{bal < 0}\ account \bowtie depositor))$

(c) $\Pi_{c\text{-name}}(\sigma_{b\text{-city} = \text{"Agra"}}\ branch \bowtie \sigma_{b\text{-city} = \text{"Agra"} \wedge bal < 0}\ account \bowtie depositor)$

(d) $\Pi_{c\text{-name}}(\sigma_{b\text{-city} = \text{"Agra"}}\ branch \bowtie (\sigma_{b\text{-city} = \text{"Agra"} \wedge bal < 0}\ account \bowtie depositor))$

Q.8 Which of the following relational query language have the same expressive power?

1. Relational Algebra.
2. Tuple relational calculus restricted to safe expression.
3. Domain relational calculus restricted to safe expression.

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

Q.9 If P and Q are predicates and e is the relational algebra expression, then which of the following equivalence is valid?

(a) $\sigma_P(\sigma_Q(e)) = \sigma_Q(\sigma_P(e))$
(b) $\sigma_P(\sigma_Q(e)) = \sigma_{P \wedge Q}(e)$
(c) $\sigma_Q(\sigma_P(e)) = \sigma_{P \wedge Q}(e)$
(d) All of the above

Q.10 Which of the following sets of operations represent a complete set of relational algebra operations?

(a) $\{\delta, \eth, \cup, -, X\}$ (b) $\{\delta, \eth, \cup, \div, X\}$
(c) $\{\delta, \eth, \cap, -, X\}$ (d) $\{\eth, \cup, \delta, -, \div\}$

Q.11 Consider the given schema:
Client (customer-name, banker-name)
Customer (customer-name, street, customer-city)
which of the following queries finds the clients of banker Atul and the city they live in?

(a) $\pi_{\text{Client.customer-name, customer-city}}(\sigma_{\text{Client.customer-name} = \text{Customer.cutomer-name}}(\sigma_{\text{Banker-name} = \text{"Atul"}}(\text{Client} \times \text{Customer})))$

(b) $\pi_{\text{customer-name, customer-city}}(\sigma_{\text{banker-name} = \text{"Atul"}}(\text{Client} \times \text{Customer}))$

(c) $\pi_{\text{Client.customer-name, customer-city}}(\sigma_{\text{banker-name} = \text{"Atul"}}(\sigma_{\text{Client.customer-name} = \text{customer.customer-name}}(\text{Client} \times \text{Customer}))$

(d) $\pi_{\text{customer-name, customer-city}}(\sigma_{\text{banker-name} = \text{"Atul"}}(\text{Client} \times \text{Customer}))$

Q.12 Translate the following tuple relational calculus expression into English language.
$\{t \mid \exists S \in \text{student}(t[S.\text{name}]) = S[S.\text{name}] \wedge \exists u \in \text{course}(u[S.\text{no}] = S[S.\text{no}] \wedge u[\text{course name}] = \text{"}CS\text{"}]]\}$

(a) Select all tuples of student name
(b) Select all tuples of student name studying "CS" course
(c) Set of all tuples of students name with same course number and student number who are studying "CS" course
(d) None of these

Q.13 The relational algebra expression equivalent to the following tuple calculate 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) $\sigma_{(A = 10)} - \sigma_{(B = 20)}(r)$

Q.14 Given the relations
Employee (name, salary, deptno) and
Department (deptno, deptname, address)
which of the following queries cannot be expressed using the basic relational algebra operation ($\sigma, \pi, \times, \bowtie, \cup, \cap, -$)?

(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 the employees of a given department

Q.15 Consider the database having the following relations:
Products (pid, pname, city-quantity, price)
Customers (cid, cname, city, discount)
Agents(aid, aname, city, percent)
Orders(orderno, month, cid, aid, pid, qty, dollors)

Which of the following returns "the customers who place order only through agent a_3"?

(a) π_{name} (Customers $\bowtie$ (π_{cid}(Customers)-$\pi_{aid}(s_{cid} = a_3$(Orders)))

(b) π_{cname}(Customers $\bowtie$ (π_{cid}(Customers)-$\pi_{cid}(s_{aid} = a_3$(Orders))))

(c) π_{cid}(Orders)-$\pi_{cid}(\sigma_{aid \neq a3}$(Orders))

(d) None of the above

Q.16 What does the following query do

{*S*. rollNo, *S*. name I student (*S*) ^ *S*.sex = '*F*' ^ (∃*d*) (department (*d*) ^ *d*. name = 'Maths' ^ *d*. deptId = *S*. deptNo)}

(a) Name of all students in Maths dept.

(b) Obtain the rollNo, name of all girl students in the Maths Dept.

(c) Obtain the name of all male students in the Maths Dept.

(d) None of the above

Q.17 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 R_2. Suppose there is no violation of the above referential integrity constraint in the corresponding relational instances r_1 and r_2.

Which one of the following relational algebra expressions would necessarily produce an empty relation?

(a) $\pi_D(r_2) - \pi_C(r_2)$ (b) $\pi_C(r_1) - \pi_D(r_2)$

(c) $\pi_D(r_1 \triangleright\!\!\triangleleft_{C \neq D} r_2)$ (d) $\pi_C(r_1 \triangleright\!\!\triangleleft_{C \neq D} r_2)$

Q.18 The following query in TRC selects:

{*S*. name | student (*S*) ^ (∃*e*) (∃*t*) (enrollment (*e*) ^ teaching (*t*) ^ *e*. courseId = *t*. courseId ^ *e*. rollNo = *S*. rollNo ^ *t*. empId = *S*. advisor}

(a) Students who have taken all courses taught by their advisor.

(b) Students who have taken courses, not taught by their advisor.

(c) Students who have taken at least one course taught by their advisor.

(d) None of the above.

Q.19 Consider the following database scheme

book (acc-no, yr-pub, title)

user (card-no, b-name, b-address)

supplier (s-name, s-adder)

borrow(acc-no, card-no, doi)

Supp (acc-no, s-name, price, dos)

What the following relational calculus represents?

{*t* | ∃*u* ∈ user (*t*[b-name] = *u*[b-name] ∧

∃*s* ∈ borrow(*u*(card-no] = *b*[card-no] ∧

(∃*s* ∈ supp(*s*[acc-no] = *b*[acc-no] ∧

s[s-name] = "narosa")) ∧

¬∃*q* ∈ supp(*q*[acc-no] = *b*[acc-no] ∧

q[s-name] = "allied"))))}

(a) Find all borrows who have issued a book supplied by "narosa" but have not issued any book supplied by "allied".

(b) Find all borrowers who have issued a book supplied by both "narosa" and allied.

(c) Find all borrowers who have issued a book supplied by 'allied' but not by 'narosa'

(d) Find atleast one borrower who have issued a book suppliedby "narosa" but have not issued any book supplied by "allied"

Q.20 Consider the sequence of operations given below on the relation

Employee (EmpNo, Name, Address, Bdate, Gender, Salary, SuperNo, DNo).

1. DEP5_EMPS ← (σDNo = 5)Employee))
2. RESULT1 ← πEmpNo(DEP5_EMPS)
3. RESULT2(EmpNo) ← πSuperNo(DEP5_EMPS)
4. RESULT ← RESULT1 ∪ RESULT2

What will the above sequence of operations performed on the given relation produce?

(a) EmpNo of the DNo 5 who work either as an employee or a supervisor.

(b) EmpNo of the employees who work in DNo 5 along with the employees of DNo 5 who work as supervisors.

(c) Emp No of the DNo 5 employees who work as supervisor.

(d) EmpNo of the employees who either work in DNo 5 or supervise an employee who works in DNo 5

Q.21 Consider the relations Branch(BranchNo, Street, City) and Property(PropertyNo, Address, City). Which of the following relational algebra expression would list all cities where there is both a branch office and at least one property?

(a) $\pi_{City}(Branch) \cup \pi_{City}(Property)$

(b) $\pi_{City}(Branch) - \pi_{City}(Property)$

(c) $\pi_{City}(Branch) \div \pi_{City}(Property)$

(d) $\pi_{City}(Branch) - (\pi_{City}(Branch) - \pi_{City}(Property)$

Q.22 Assume table R has two attributes A and B. Similarly table S has two attributes A and B. Which of the following relational algebra expressions are not equivalent to $R \cap S$?

(a) $R \bowtie S$

(b) $(R \cup S) - ((R - S) \cup (S - R))$

(c) $S - (S - R)$

(d) $R - (S - R)$

Q.23 Find the number of tuples returned by the following query? [ρ is used to rename]

$$\pi_{AD}(R \times S) - \rho_{A \leftarrow B}(\pi_{BD}(R \bowtie_{B=C} S))$$

(a) 4 (b) 5

(c) 6 (d) 7

Q.24 Member (Mid, Name, Designation, Age)
Books (Bid, Btitle, Bauthor, Bprice)
Reserves (Mid, Bid, Date)
Which of the following option will retrieve the titles of books reserved by lecturers?

(a) $\Pi_{Name,Btitile}((Member \bowtie_{(Member.Mid = reserve.Mid)\ AND\ (Designation = 'Lecturer')} Reserve) \bowtie_{res.bid = Books.Bid} Books)$

(b) $\Pi_{Name,Btitile}((Member_{(Member.Mid = reserve.Mid)\ AND\ (Designation = 'Lecturer')} Reserve) \bowtie_{res.lect.bid = Books.Bid} Books)$

(c) $\Pi_{Name'Btitile}((Member \bowtie_{(Member.Mid = reserve.Mid)\ AND\ (Designation = 'Lecturer')} Reserve) \bowtie_{res.tect.bid = Books.Bid} Books)$

(d) None of these

Q.25 Let R and S be two relations with the following schema

$R(\underline{P}\ \underline{Q}\ R_1\ R_2\ R_3)$

$S(\underline{P}\ \underline{Q}\ S_1\ S_2)$

Where $\{P, Q\}$ is the key for both schemas. Which of the following queries are equivalent?

I. $\Pi_P(R \bowtie S)$

II. $\Pi_P(R) \bowtie \Pi_P(S)$

III. $\Pi_P(\Pi_{P,Q}(R) \cap \Pi_{P,Q} S))$

IV. $\Pi_P(\Pi_{P,Q}(R) - (\Pi_{P,Q}(R) - \Pi_{P,Q}(S)))$

(a) Only I and II (b) Only I and III

(c) Only I, II and III (d) Only I, III and IV

Q.26 Consider the following relations:

$R_1(A\ B\ C)$ and $R_2(A\ D\ E)$

R_1 has 1000 records and R_2 has 2000 records. The non-Null attribute 'A' in R_2 is referencing attribute 'A' in R_1. Let X be minimum number of records in $R_1 \bowtie R_2$ and Y be the maximum number of records in $R_1 \bowtie R_2$. The sum of $(X + Y)$ is _____.

Q.27 Consider a join (relation alzebra) between relation $r(R)$ with 'n' blocks with each block contain 'R_n' records and $s(S)$ with 'm' blocks with each block contain 'R_m' records using block nested loop over nested loop join if at any time only 2 block present in main memory. Which of the following is true?

(a) If $r(R)$ is outer loop then access cost is $n + n \times m$.

(b) If $r(R)$ is outer loop then access cost is $m + R_m \times n$.

(c) If $s(S)$ is outer loop then access cost is $n + R_n \times m$.

(d) None of the above

Q.28 Consider the relations $P(\underline{A}, B)$ and $Q(A, \underline{B})$ where P has foreign key referencing Q via B and Q has foreign key referencing P via A. Which of the following is guaranteed to produce fewer than or at most the same number of tuples as any of the other tables?

(a) $P \bowtie \pi_B(Q)$ (b) $\pi_A(P) \bowtie Q$

(c) $P \bowtie Q$ (d) None of these

Q.29 Consider the following relations R and S:

R

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

S

B	C	D
2	3	10
2	3	11
6	7	12
7	8	10

The number of tuples in $R \bowtie S$ [right outer join] are _____.

Q.30 Consider a selection query: $\sigma_{A \le 15}(R)$, where R is a relation with 400 tuples. Assume that the attribute values (integers) for A among the tuples are uniformly distributed in the interval [1, 25]. How many tuples would the given selection query returns?

Q.31 Consider the following relation schema:

Student (Sid, Sname)

Course (Cid, Cname, Dname)

Takes (Sid, Cid, Dname)

"Find the name of all the students who have taken all the course of the department '*CS*".

Which of the following will represents the above query?

(a) $\{t \mid \exists r \in \text{Student}\ (r[\text{Sid}]=t[\text{Sid}]) \wedge (\forall u \in \text{Course}\ (u[\text{Dname}]=\text{"}CS\text{"} \Rightarrow \exists s \in \text{Takes}\ (t[\text{Sid}] = s[\text{Sid}] \wedge s[\text{cid}] = u[\text{cid}]))\}$

(b) $\{t \mid \exists r \in \text{Student}\ (r[\text{Sid}]=t[\text{Sid}]) \wedge (\forall u \in \text{Course}\ (u[\text{Dname}] \neq \text{"}CS\text{"} \vee \exists s \in \text{Takes}\ (t[\text{Sid}] = s[\text{Sid}] \wedge s[\text{cid}] = u[\text{cid}]))\}$

(c) Both (a) and (b)

(d) None of the above

Q.32 Consider two relations $R(A_1, A_2, A_3, \ldots, A_m)$ and $S(A_1, A_2, A_3, \ldots, A_n)$ where $m > n$. Find the number of attributes that appears in R / S.

(a) $\left\lfloor \frac{n}{m} \right\rfloor$ (b) $\left\lfloor \frac{m}{n} \right\rfloor$

(c) $m - n$ (d) $m + n$

Q.33 Assume $R(A, B)$ and $S(C, D)$ relations have the following instances

R

A	B
1	2
2	1
3	3

S

C	D
1	2
3	4
3	5

Find number of tuples returned by the following query (r is used to rename the attribute)

$\pi_{AD}(R \times S) - \rho_{A \leftarrow B}\left(\pi_{BD}\left(R \bowtie_{B=C} S\right)\right)$

Q.34 Consider the relation $P(A, B, \underline{C})$, $Q(\underline{C}, D, \underline{E})$ and $R(\underline{E}, F)$ having tuples 200, 300 and 100 respectively. Estimate the number of tuples in relation $P \bowtie Q \bowtie R$.

Q.35 Consider two relations enrolled and course as shown below:

Enrolled

Sid	Cid	Fees
S1	C1	10
S1	C2	20
S2	C3	30
S3	C4	40

Course

Cid	Cname	Dept
C1	ALGO	CS
C2	DS	CS
C3	TOC	CS
C4	THERMO	ME

$\pi_{\text{Sid,Cid}}(\text{Enrolled}) / \pi_{\text{Cid}}(\sigma_{\text{Dept} = \text{'EE'}}(\text{Course}))$.

If above relational algebra query executes over above data base table, then how many tuples are there in the result of query?

■■■■

Answers Relational Algebra

1.	(d)	2.	(b)	3.	(d)	4.	(d)	5.	(d)	6.	(d)	7.	(b)	8.	(d)	9.	(d)
10.	(a)	11.	(a)	12.	(b)	13.	(c)	14.	(c)	15.	(c)	16.	(b)	17.	(b)	18.	(c)
19.	(a)	20.	(d)	21.	(d)	22.	(d)	23.	(c)	24.	(a)	25.	(c)	27.	(a)	28.	(c)
31.	(c)	32.	(c)														

Explanations Relational Algebra

2. **(b)**
Loans (br_name, loan_no, amount)
For the query "find the loan number for each loan of an amount greater than 2000".
we have to check for each loan to find a loan number which is greater than 20,000.
clearly, (b) $[t \mid \exists\ s \in$ loans $(t$ [loan_no.] = S[loan_no.] $\wedge$ [amount] > 20000$\}$

3. **(d)**
$E_1 \bowtie_\theta E_2 = E \bowtie_{\theta 1} E_1$ is not equivalent because θ join is not commulative.
Hence the entries in $E_1 \bowtie_\theta E_2$ and $E_2 \bowtie_{\theta 1} E_1$ may not be same.

4. **(d)**
Stud Info(stud Id, name, sex)
Enrol (Stud Id, course Id)
The relational algebra expression
$\pi_{\text{course Id}}((\pi_{\text{stud Id}}(\sigma_{\text{sex = "female"}}(\text{stud Info})) \times \pi_{\text{course Id}}(\text{enrol}) - \text{enrol})$
represents courses in which a proper subset of female students are enrolled.
Hence, option (d) is correct.

5. **(d)**

a	c
1	2
2	3

As we can see clearly both entries of $(a < c)$. also for selecting the entries we have to project on R.
Hence (d) $\sigma_{a<c}(\Pi_{a,c} R)$ is the correct option.

7. **(b)**
Branch: (b_name, b_city, assets)
Account: (a_num, b_name, bal)
depositor: (c_name, a_number)
$\pi_{c\_name}(\sigma_{b\_city = \text{'Agra'} \wedge bal < 0}(\text{branch} \bowtie (\text{account} \bowtie \text{depositor})))$
(a) $\pi_{c\_name}(\sigma_{bal<0}(\sigma_{b\_city = \text{'Agra'}}$ branch $\bowtie$ account $\bowtie$ depositor))
This join candition will increase the number of results / tuples and at the end, condition is applied which is inefficient.
(b) $\pi_{c\_name}(\sigma_{b\_city = \text{'Agra'}}$ branch $\bowtie (\sigma_{bal<0}$ account $\bowtie$ depositor))
$\sigma_{b\_city = \text{'Agra'}}$ branch gives the tuples with b_city = 'Agra' ...(2)
$\sigma_{bal<0}$ account $\bowtie$ depositor finds those acc_num balance less than 0 ...(1)
The join of (1) and (2) gives less number of tuples comparatively option (c) and (d) will also give more number of tuples because of the join conditions and the join operations of tables.
So, option (b) is correct.

9. **(d)**
All of the above are valid the conditions basically are evaluating $P \wedge Q$.

10. **(a)**
$\{\sigma, \pi, \cup, -, X\}$ these contribute a complete set of relational algebra operation.

11. **(a)**
$\sigma_{\text{banker\_name= 'Atul'}}(\text{client} \times \text{Customer})$ ...(1)
It finds the combinations of all clients and customer whose banker name is Atul.
$\sigma_{\text{client.customer\_name = Customer.customer\_name}}((1))$
It finds those tuples where both tables have same customer name with banker_name as Atul and finally, the π conditions retrieves their names and customer-city of clients of banker 'Atul'.

12. **(b)**
The tuple relationa calculas evaluates all those students having course_name as 'CS'.

13. **(c)**
Option (c) is valid as it calculates the **intersection** of different tables having $A = 10$ and $B = 20$ separately.

14. **(c)**
The sum operation can't be performed with the basic algebra operation.

15. (c)

$\pi_{cid}(orders) - \pi_{cid}(\sigma_{aid \neq a3}(orders))$

The above query evaluates those customers who ordered through agent a_3.

16. (b)

Query gives the roll no. and name of those students who are female (sex=f) and who are in maths department (d.name= 'math').

17. (b)

As C is refering to R_2 and D is primary key of R_2, $\pi_C(r_1) - \pi_D(r_2)$ will give empty relation or empty table as number of values in C column of table r_1 will always refer to of respective values in D column of r_2.

18. (c)

From the options, option (c) is giving the best explanation of the query that it gives names of all students who have taken at least 1 course under their advisor.

19. (a)

The query gives all those names who have issued/borrow a book from 'Narosa' but not 'Allied'.

20. (d)

It gives the union of all the employees who are working in D.No. 5 or are supervisor for the employee who are working in D.No.5.

21. (d)

$\pi_{city}(branch) - \pi_{city}(property)$ ...(1)

It gives those city names where there is no property.

$\pi_{city}(branch) - ((1))$

It gives those city names having branch office and at least one property as cities with property have been eliminated.

22. (d)

$R-(S-R)$ is equivalent to R.

Except (d), remaining all relational algebra expressions are equivalent.

23. (c)

$\pi_{AD}(R \times S) - \rho_{A \leftarrow B}(\pi_{BD}(R \bowtie_{B=C} S))$

A	D
a	b
a	d
a	e
b	b
b	d
b	e
c	b
c	d
c	e

$-\rho_{A \leftarrow B}$

B	D
a	b
c	d
c	e

$=$

A	D
a	d
a	e
b	b
b	d
b	e
c	b

$\therefore$ 6 tuples are returned by the given query.

24. (a)

Option (a) is giving those book names and member names combination who have reserved the books and who are lecturers.

25. (c)

R and S are two relations

$R(P\ Q\ R_1\ R_2\ R_3)$

$S(P\ Q\ S_1\ S_2)$

(i) $\Pi_P(R \bowtie S)$: The query join the relation R and S then project the column P only.

(ii) $\Pi_P(R) \bowtie \Pi_P(S)$: The query project the column P from R and S and then join these projected columns so this produces same result in (i).

(iii) $\Pi_P(\Pi_{P,Q}(R) \cap \Pi_{P,Q}(S))$: The given query project separately the column P and Q in both relations R and S. The intersection produces only these column which are common in both and finally projection produces the result the column P so this query is also equivalent to (i) and (ii).

26. (4000)

Since 'A' in R_2 is not key, hence all value of A may or may not be unique. Hence every entry under A in R_2 will match with A in R_1.

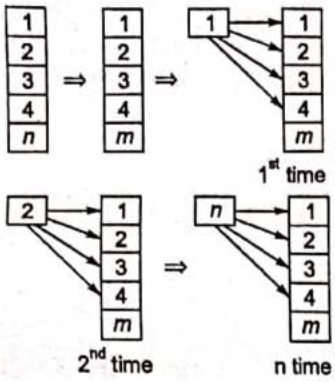
Hence maximum is 2000. But A in R_2 is foreign key referencing A in R_1.

Therefore minimum also 2000.

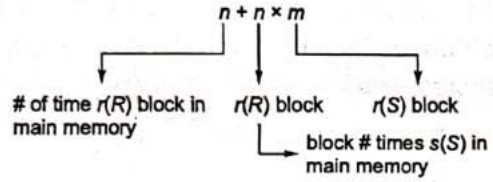
$\therefore$ Sum is 2000 + 2000 = 4000

27. (a)

Block nested loop join.



i.e. If $r(R)$ is outer



28. (c)

$P \bowtie Q$ is guaranteed to produce fewer than or at most the same number of tuples as any of the other tables.

29. (4)

$R \bowtie\!\!\!\subset S$ [right outer join of R and S]

A	B	C	D
1	2	3	10
1	2	3	11
6	7	8	10
NULL	6	7	12

$\therefore$ Number of tuples = 4

30. (240)

R contain 400 tuples and values for A are uniformly distributed in the interval [1, 25]

Number of values in [1, 25] = 25

Since numbers are uniformly distributed then a

$$\text{number repeats} = \frac{400}{25} = 16$$

$\therefore$ Each value in domain repeats for 16 times.

So number of tuples $\leq 15 = 15$

Hence, $\sigma_{A \leq 15}(R) = 15 \times 16 = 240$ tuples.

31. (c)

"Find the name of all the students who have taken all the course of the department '*CS*'".

$\{t \mid \exists r \in \text{Student}\ (r[\text{Sid}] = t[\text{Sid}]) \wedge (\forall u \in \text{Course}\ (u[\text{Dname}] = \text{"}CS\text{"} \Rightarrow \exists s \in \text{Takes}\ (t[\text{Sid}] = s[\text{Sid}] \wedge s[\text{cid}] = u[\text{cid}]))\}$

Since $P \Rightarrow Q \equiv \sim P \vee Q$ therefore $\{t \mid \exists r \in \text{Student}\ (r[\text{Sid}] = t[\text{Sid}]) \wedge (\forall u \in \text{Course}\ (u[\text{Dname}] \neq \text{"}CS\text{"} \vee \exists s \in \text{Takes}\ (t[\text{Sid}] = s[\text{Sid}] \wedge s[\text{cid}] = u[\text{cid}]))\}$.

32. (c)

Number of attributes in R/S is $m - n$.

33. (6)

$\pi_{AD}(R \times S)$

A	D
1	2 ✓
1	4
1	5
2	2
2	4
2	5
3	2
3	4 ✓
3	5 ✓

$- \rho_{A \leftarrow B}$

B	D
1	2
3	4
3	5

$=$

A	D
1	4
1	5
2	2
2	4
2	5
3	2

6 tuples are returned by the query.

34. (100)

$P \bowtie Q \Rightarrow$ Common attribute = C which is key for both P and Q.

$R_1 = P \bowtie Q \Rightarrow \min(200, 300) = 200$ tuples

$R_1 \bowtie R \Rightarrow$ Common attribute = E which is key for both R_1 and R.

$R_1 \bowtie R \Rightarrow \min(200, 100) = 100$ tuples

So, $|P \bowtie Q \bowtie R| = 100$ tuples

35. (3)

$|\pi_{\text{Cid}}(\sigma_{\text{dept} = \text{'EE'}}(\text{Course})| = 0$.

$\pi_{\text{Sid,Cid}}(\text{Enrolled}) / \pi_{\text{Cid}}(\text{empty tuples in Dept})$

All the distinct Sid will be in the result.

■■■■

4 CHAPTER

SQL

Q.1 Consider the following queries:

S_1 : SELECT e.name
FROM emp e
WHERE e.sal = ANY
(SELECT sal FROM emp
WHERE e.name = 'x');

S_2 : SELECT e.name
FROM emp e
WHERE e.sal IN
(SELECT sal FROM emp
WHERE e.name = 'x');

S_3 : SELECT e.name
FROM emp e
WHERE e.sal = ALL
(SELECT sal FROM emp
WHERE e.name = 'x');

Which one of the following is true?

(a) S_1, S_2 are same (b) S_2, S_3 are same
(c) S_1, S_2, S_3 are same (d) None of these

Q.2 Given relations $R(w, x)$ and $S(y, z)$ the result of SELECT DISTINCT w, x FROM R, S; is guranteed to be same as R, if

(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

Q.3 A table $T1$ in a relational database has the following rows and columns:

Roll No	Marks
1	110
2	20
3	30
4	Null

The following sequence of SQL statements was successfully 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.4 Consider the SQL query: SELECT DISTINCT a_1, a_2,..., a_n FROM r_1, r_2,..., r_m WHERE P. For an arbitrary predicate P, this query is equivalent to which of the following relational algebra expressions?

(a) $\pi_{a_1, a_2 \ldots a_n}\ \sigma_p\ (r_1 \times r_2 \times \ldots \times r_m)$

(b) $\pi_{a_1, a_2 \ldots a_n}\ \sigma_p\ (r_1 \bowtie r_2 \times \ldots \bowtie r_m)$

(c) $\pi_{a_1, a_2 \ldots a_n}\ \sigma_p\ (r_1 \cup r_2 \cup \ldots \cup r_m)$

(d) $\pi_{a_1, a_2 \ldots a_n}\ \sigma_p\ (r_1 \cap r_2 \cap \ldots \cap r_m)$

Q.5 Consider the following queries:

Q_1 : SELECT max (sal) FROM emp
GROUP BY dept no HAVING dept no <>10;

Q_2 : SELECT max (sal) FROM emp
WHERE dept no <> 10 GROUP BY dept no;

Which of the following is false about queries?

(a) Both queries can be used for the required result
(b) Q_1 is faster than Q_2
(c) Q_2 is faster than Q_1
(d) None of these

Q.6 Consider the following relation schema pertaining to suppliers_parts database
S(S#, SNAME), P (P#, COLOR) and SP denotes the product of S and P. What does the following SQL query produces?

SELECT DISTINCT S. NAME
FROM S
WHERE S. S # IN (SELECT SP. S# FROM SP
WHERE SP. P# IN (SELECT P.P# FROM P WHERE
P. COLOR = 'RED')),

(a) Get supplier names for suppliers who supply only RED parts
(b) Get supplier names for suppliers who supply atleast one red part
(c) Get supplier name for suppliers who do not supply red parts
(d) None of the above

Q.7 Consider two tables in a relational database with columns and rows as follows:

Table : Student

Roll_no	Name	Dept_id
1	ABC	1
2	DEF	2
3	GHI	3
4	JKL	4

Table : Department

Dept_id	Dept_name
1	A
2	B
3	C

Roll_no is the primary key of the student table, Dept_id is the primary key of the department table and student. Dept_id is a foreign key from Department. Dept_id

What will happen if we try to execute the following two SQL statements?

1. update student set Dept_id = Null where Roll_no =1
2. update Department set Dept_id = Null where Dept_id = 1

(a) Both 1 and 2 will fail
(b) 1 will fail 2 will succeed
(c) 1 will succeed but 2 will fail
(d) Both 1 and 2 will succeed

Q.8 Let $R(A, B, C)$ and $S(B, C, D)$ be the relations as given below:

R

A	B	C
a1	b1	c2
a2	b4	c1
a3	b3	c3
a2	b3	c5
a1	b1	c2

S

B	C	D
b1	c1	d2
b2	c4	d1
b3	c3	d4

What is the result of following query?
SELECT R.A FROM R WHERE R.B = NOT ALL (SELECT S.B FROM S WHERE R.C. = S.C)

(a)

A
a1
a3
a2
a1

(b)

A
a1
a3
a2

(c)

A
a3
a2

(d)

A
a1
a3
a2
a3

Q.9 A company has a storeroom, which consists of few employees who order products from different shops. Following is the snapshot of the "product-order" table maintained at company's database.

Product-Order

Product-no	Employee-no	Shop-no
P1	E6	S2
P2	E2	S1
P3	E7	S4
P4	E5	S3
P5	E2	S3
P6	E7	S2

Consider the following query:
SELECT A.product_no, A.employee_no
FROM product_order *A*, product_order *B*
WHERE A.employee_no = *B*.employee_no
and *A*.shop_no < > *B*.shop_no;
The number of tuples contained in the output will be

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

Q.10 There are six tables describing a company, describing employees, departments, buildings, which department(s) an employee works in (and a percentage of the time for each), department managers (possibly more than one per department), and in which building an employee works (an employee may have more than one office). The primary key of each table is the attribute(s) in capitals. Other attributes are not necessarily unique.

EMP (EID, EName, Salary, Start_Date, End_Date)
IN_DEPT (EID, DID, Percent_Time)
BUILDING (BID, Bname, Address)
IN_BUILDING (EID, BID)
DEPT (DID, Dname, Annual_Budget)
MANAGES_DEPT (EID, DID)

Which of the following queries finds the name of Departments where no employees work?

(a) SELECT Dname
FROM DEPT
WHERE DID IN(SELECT I.DID FROM
IN_DEPT I
GROUP BY I.DID HAVING COUNT(*) = 0)

(b) SELECT Dname
FROM DEPT D, IN_DEPT I, EMP E
WHERE I.EID = E.EID and
D.DID=I.DID and Count(E.EID) = 0

(c) SELECT Dname
FROM DEPT , IN_DEPT
WHERE IN_DEPT.DID = DEPT.DID GROUP
BY DID
HAVING COUNT (EID) = 0

(d) Both (b) and (c)

Q.11 Consider the following relations:
Student (snum, sname, major, level, age)
Class (name, meets_at, room, fid)
Enrolled (snum, cname)
Faculty (fid, fname, deptid)

What does the following query find?
SELECT C.name
FROM Class C
WHERE C.room = 'R128'
or C.name in (SELECT E.Cname) FROM Enrolled *E* GROUP BY *E*.Cname HAVING count(*) >= 5)

(a) Finds the names of all classes that either meet in room R128 and have five or more students enrolled
(b) Finds the names of all classes that either meet in room R128 or have five or more students enrolled
(c) Finds the name of atmost one class that either meet in room R128 or have five or more students enrolled
(d) Finds the names of all classes that meet in room R128 that has five or more students enrolled in it

Q.12 Which of the following SQL query is illegal.
(a) Select max (count(*)) from *R* group by *A*;
(b) Select count(*), max (*B*) from R group by *A*;
(c) Select count(*) from *R* group by *A*;
(d) Select count(*) from *R*;

Q.13 Consider the following SQL queries
Q_1: Select *x* from *R r*1
where EXISTS (select* from *R* where $x = r1.y$);
Q_2: Select *x* from *R*
where *y* = ANY (select *x* from *R*);

Which of the following is correct?
(a) Q_1 and Q_2 returns same result.
(b) Q_1 can return less rows than returned by Q_2.
(c) Q_1 can return more rows than returned by Q_2.
(d) Q_1 and Q_2 results have no relation.

Q.14 Assume *R*(*A*, *B*) and *S*(*C*, *D*) relations have the following instances.

R

A	B
a	a
b	a
c	c

S

C	D
a	b
c	d
c	e

Find the number of tuples returned by the following query ?
(Select *B* from *R*)
Union all
(Select distinct *C* from *S*)
(a) 2 (b) 3
(c) 4 (d) 5

Q.15 Find the equivalent SQL query for the following relation Algebra query

$$\Pi_{A,B}(\sigma_{C=D}(x, y))$$

(a) Select *A*, *B* from *x*, *y* where *C* = *D*
(b) Select distict *A*, *B* from *x*, *y* where *C* = *D*
(c) Select * from *x*, *y* where *C* = *D*
(d) None of these

Q.16 Which of the following queries will give the names of employee who is earning maximum salary?

(a) Select name from emp
where sal = (select max (sal) from emp);
(b) Select name from emp
where sal > = all (select sal from emp);

(c) Both (a) and (b) are correct queries but processing is too high in B than in A

(d) Both (a) and (b) are correct queries but processing is too high in A than in B

Q.17 Consider the statements

$S1$: "Delete" is used to delete the table from database.

$S2$: "Truncate table" is used to delete all the data but not table.

$S3$: "Drop table" is used to delete the data as well as table.

Which of the above statement(s) is/are true?

(a) $S1$ and $S2$ (b) $S2$ and $S3$

(c) $S1$ and $S3$ (d) $S1$, $S2$ and $S3$

Q.18 Consider the following statements with queries.

$S1$: [Select * from R_1, R_2] is equal to $[R_1 \times R_2]$

$S2$: $[R_1 \bowtie_\theta R_2]$ is equal to $[\sigma_\theta(R_1 \times R_2)]$

$S3$: $[(R - T) \bowtie S]$ is equal to $(R - T) \bowtie S$

Identify the valid statements?

(a) Only $S1$ (b) $S1$, $S2$

(c) $S2$, $S3$ (d) $S1$, $S2$, $S3$

Q.19 Consider the following query languages.

L_1: Relational Algebra.

L_2: Domain relational calculus.

L_3: Tuple relational calculus.

L_4: SQL without the aggregate operators or recursion.

Find the equally expressive query languages? (All queries possible to state in X language iff all those queries can be stated in other Y language represents X and Y are equally expressive query languages).

(a) L_1, L_2, L_3 (b) L_2, L_3, L_4

(c) L_1, L_2, L_4 (d) L_1, L_2, L_3, L_4

Q.20 The standard SQL "ORDER BY" clause

(a) can be used to order the columns used in select.

(b) may only be used when "GROUP BY" is used.

(c) can be used any place of query to order the tuples.

(d) can only be used as the last clause of a query to order the tuples.

Q.21 Which of the following is correct?

(a) An SQL query automatically eliminates duplicates

(b) AN SQL query will not work if there are no indexes on the relations

(c) SQL permits attribute names to be repeated in the same relation

(d) None of the above

Q.22 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

Q.23 Consider the set of relations given below and the SQL query that follows:

Students: (Roll_number, Name, date_of_birth)
Courses: (Course_number, Course_name, Instructor)
Grades: (Roll_number, Course_number, Grade)

```
SELECT DISTINCT Name
FROM Students, Courses, Grades
WHERE Students. Roll_number = Grades.Roll_number
AND Course.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

Q.24 Consider the relations:

Supplier (S_no, S_name, city) and

Item (Item_no, Item_name, Item_Brand, S_no)

Which of the following SQL statement gives suppliers names who have submitted at least one 'Nestle' brand item?

(a) SELECT S_name FROM Supplier S, Item I WHERE I. Item_Brand = 'Nestle'

(b) SELECT S_name FROM Supplier S WHERE S_no IN (SELECT S_no From Item WHERE Item_Brand = 'Nestle')

(c) SELECT S_name FROM Supplier S WHERE Item_Brand = 'Nestle'

(d) SELECT S_name FROM Supplier WHERE S_no EXISTS (SELECT S_no FROM Supplier, Item WHERE Item_Brand = 'Nestle')

Q.25 There are two relations X and Y. Relation X has arity 1 and cardinality 2, relation Y has arity 3 and cardinality 4. Indicate the result of the SQL statement SELECT COUNT (*) FROM X, Y.

(a) 4

(b) 6

(c) 8

(d) Can not be calculated from given information

Q.26 Consider the table employee (empId, name, department, salary) and the two queries Q_1, Q_2 below. Assuming that department 5 has more than one employee, and we want to find the employees who get higher salary than anyone in the department 5, which one of the statements is TRUE for any arbitrary employee table?

Q_1 : Select e.empId
From employee e
Where not exists
(Select * From employee s Where s. department = "5" and s. salary > = e.salary)

Q_2 : Select e.empId
From employee e
Where e.salary > Any
(Select distinct salary From employee s
Where s. department ="5")

(a) Q_1 is the correct query.

(b) Q_2 is the correct query.

(c) Both Q_1 and Q_2 produce the same answer

(d) Neither Q_1 nor Q_2 is the correct query

Q.27 Consider a database with three relation instance shown below. The primary keys for the Drivers and Cars relation are did and cid respectively and the records are stored in ascending order of these primary keys as given in the tables. No indexing is available in the database.

D : Drivers relation

did	dname	rating	age
22	Karthikeyan	7	25
29	Salman	1	33
31	Boris	8	55
32	Amoldt	8	25
58	Schumacher	10	35
64	Sachin	7	35
71	Senna	10	16
74	Sachin	9	35
85	Rahul	3	25
95	Ralph	3	53

R : Reserves relation

did	cid	day
22	101	10/10/06
22	102	10/10/06
22	103	08/10/06
22	104	7/10/06
31	102	10/11/06
31	103	06/11/06
31	104	12/11/06
64	101	05/09/06
64	102	08/09/06
74	103	08/09/06

C : Cars relation

cid	cname	color
101	Renault	blue
102	Renault	red
103	Ferrari	green
104	Jaguar	red

What is the output of the following SQL query?

```
select D. daname
        from Drivers D
        where D. did in (
        select R. did
        from Cars C, Reserves R
where  R.cid = C.cid and C. color = 'red'
        intersect
        select R.did
        from Cars C, Reserves R
where R.cid = C.cid and C. colour = 'green' )
```

(a) Karthikeyan, Boris
(b) Sachin, Salman
(c) Karthikeyan, Boris, Sachin
(d) Schumacher, Senna

Q.28 An athletics meeting involves several competitors who participate in a number of events. The database is intended to record who is to take part in which event and to record the outcome of each event. As results become available the winner attribute will be updated with the cid of the appropriate competitor.

Competitor (<u>cid</u>, name, nationality)
Even (<u>eid</u>, description)
Competes (<u>cid</u>, <u>eid</u>)

Competitor

cid	name	nationality
01	Pat	British
02	Hilary	British
03	Sven	Swedish
04	Pierre	French

Event

eid	description
01	running
02	jumping
03	throwing

Competes

cid	eid
01	01
02	01
03	02
04	02
04	03

Identify the result of the following SQL statement.

```
SELECT eid FROM Competes, Competitor
WHERE Competes. cid = Competitor. cid
AND nationality = 'Swedish';
```

(a) 01 (b) 02
(c) 04 (d) 03

Q.29 Table GATE2016 has 16 records. It has a non-null 'Marks' column which is also unique. Consider the following SQL query.

```
SELECT COUNT(*)
FROM GATE2016
WHERE MARKS > ANY (SELECT MARKS FROM
GATE2016).
```

The value printed by above query is _____.

Q.30 In SQL the statement **Select * from *R*, *S*** is equivalent to

(a) Select * From *R* natural join *S*
(b) Select * From *R* cross join *S*
(c) Select * From *R* union join *S*
(d) Select * From *R* inner join *S*

Q.31 Consider the two tables in a relational database with columns and rows as follows:

Table : Employee

ID	Name	Did
1	Suresh	1
2	Shiva	1
3	Mallesham	2
4	Aditya	3

Table : Department

Did	Dname
1	CS
2	EC
3	ME

ID is the primary key of Employee table, Did is the primary key of the Department table. Employee.Did is a foreign key refers Department.Did

What will happen if we try to execute the following two SQL statements?

(*i*) Update Employee set Did=Null where ID=1;
(*ii*) Update Department set Did=Null where Did=1;

(a) Both (*i*) and (*ii*) will fail
(b) (*i*) will fail but (*ii*) will succeed
(c) (*i*) will succeed but (*ii*) will fail
(d) Both (*i*) and (*ii*) will succeed

Q.32 A database table *T*1 has 4000 records and occupies 50 disk blocks. Another table *T*2 has 300 records and occupies 20 disk blocks. These two tables have to be joined as per a specified join condition that needs to be evaluated for every pair of records from these two tables. The memory buffer 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. If Nested-loop join algorithm is employed to perform the join, with the most appropriate choice of table to be used in outer loop, the number of block accesses required for reading the data are _____.

Q.33 Consider the following SQL statements with its number.

SELECT	1
FROM	2
WHERE	3
GROUP BY	4
HAVING	5
ORDER BY	6

What is correct order (number sequence) for evaluating an SQL statements? (where order is 6 digit number)

Q.34 Consider the following schema.

$R(\underline{C}, D), Q(\underline{B}, C), P(\underline{A}, B)$

C is foreign key in Q referencing $R(C)$ on delete cascade, B is foreign key in P referencing $Q(B)$ on delete set null. Suppose current content of P, Q, R as follows.

P

A	B
a	a
b	b

Q

B	C
a	a
b	a

R

C	D
a	a
b	a

After executing the following query, what are the tuples in P?

delete from R

(a) (a, NULL) and (b, b)
(b) (a, NULL) and (b, NULL)
(c) (b, b) only
(d) P will not be changed

Q.35 Consider the following relations:

Bank(bname, city)
Travel(pname, city)

```
SELECT T1.pname
FROM Travel T1
WHERE NOT EXISTS (SELECT B.city
                  From Bank B
                  WHERE B.bname = 'SBI'
                  EXCEPT
                  SELECT T2.city
                  FROM Travel T2
WHERE T1.pname = T2.pname)
```

The above query finds name of the persons.

(a) Who have not travelled in any city where SBI is located
(b) Who have not travelled in all city where SBI is located
(c) Who have travelled in all city where SBI is located
(d) Who have travelled in any city where SBI is located

Q.36 Consider the following table:

Gate 2016

Test id	Student Name	GATE Score
1	Nagender	100
2	Amjad	300
3	Sachin	300
4	Anand	200
5	Mukesh	150
6	Sushil	250

How many tuples result after executing the query:

```
SELECT Student Name, SUM (Gate Score) from
Gate 2016 group by Student Name
Having sum (Gate Score) < 500;
```

Q.37 Consider the relation and query:

Project (P.no, Pname, budget, city)

```
SELECT Pname
FROM project
where NOT(budget >= ANY ( SELECT budget
                          FROM project
                          where city=='KANPUR'))
```

Which of the following output produced by above query?

(a) Name of projects whoes budget is less than every project in Kanpur city.
(b) Name of projects whoes budget is greater than every project in Kanpur city.
(c) Name of projects whoes budget is less than equal to project in Kanpur city.
(d) Name of projects whoes budget is greater than equal to every project in Kanpur city.

■■■■

Answers SQL

1.	(a)	2.	(a)	3.	(c)	4.	(a)	5.	(b)	6.	(b)	7.	(c)	8.	(b)	9.	(b)
10.	(c)	11.	(b)	12.	(a)	13.	(a)	14.	(d)	15.	(b)	16.	(c)	17.	(b)	18.	(d)
19.	(d)	20.	(d)	21.	(d)	22.	(d)	23.	(c)	24.	(b)	25.	(c)	26.	(b)	27.	(a)
28.	(b)	30.	(b)	31.	(c)	34.	(b)	35.	(c)	37.	(a)						

Explanations SQL

1. (a)
S_1 and S_2 are same as 'ANY' and 'IN' conditions will give the same result.

2. (a)
If R contains duplicate values then the query will not select those duplicate values, i.e. $R \times S \subset R$ and $R \times S \neq R$. S may or may not have duplicates.

3. (c)
The UPDATE command will leave the table $T1$ in the following state

Roll No.	Marks
1	15
2	25
3	35
4	Null

All aggregate functions except COUNT (*) ignore NULL values in the column. In our case the average will be calculated based only on the rows in the table where a valid value is stored in the marks column.

Hence, AVG (marks) $= \frac{15+25+35}{3} = \frac{75}{3} = 25$

4. (a)
Cross product of the relations are taken which are in the FROM clause. Predicate P selects the tuples that satisfy it. Finally DISTINCT removes the duplicate tuples.

5. (b)
Q_1 is slower than Q_2 because Q_1 groups the department no. without filtering. Whereas Q_2 filters the data first that is it removes all those tuples where department No. = 10 and then groups them.

6. (b)
The inner most query returns $P\#$ of those parts whose color is red. The outer query returns all those suppliers who supply atleast one red part "Atleast" is used because of IN operator.

7. (c)
Query-1 will succeed as dept_id of student table is foreign key. But in query-2, dept_id is of department table which is primary key and which can't be set to NULL.

8. (b)
Select $b.B$ from S where $R.C = S.C$

It will return $\begin{array}{|c|}\hline b_1 \\ \hline b_3 \\ \hline \end{array}$ as c_1 and c_3 matches in both R and S.

The total query will give, $\begin{array}{|c|}\hline a_1 \\ \hline a_3 \\ \hline a_2 \\ \hline \end{array}$

9. (b)
The output Table will be

Product_no.	employ_no.
P_2	E_2
P_3	E_7
P_5	E_2
P_6	E_7

10. (c)
Option (c) gives the result of the query stated as dept DID should match IN_DEPT DID.

11. (b)
The query gives the names of those classes whose room number is 'R128' or have 5 or more students enrolled.

12. (a)
Option (a) query is illegal due to max (count(*)).

13. (a)
Q_1 and Q_2 return the same result
∴ Both queries are same

14. (d)
(1) Select B from $R \Rightarrow$ returns 3 tuples.
(2) Select distinct C from $S \Rightarrow$ returns 2 tuples
(1) Union all (2) $\Rightarrow$ returns 3 + 2 = 5 tuples (keep duplicates)

15. (b)
$\pi_{A,B}(\sigma_{C=D}(x, y))$
It finds A and B values from x and y for which $C = D$. Relational algebra also eliminates all duplicate values.

16. (c)
Both (a) and (b) will give the name of employee who is earning maximum salary.
(a) Select name from emp
where sal = (select max(sal) from emp);
in this nested query the inner query will go through all n tuples of emp database table, and will output 1 record. For this 1 record the outer query will output the name. So the processing time is $0(n)$.
(b) Select name from emp
where sal > = all (select sal from emp);
in this nested query for every tuple of outer query, its salary attribute will be checked with all tuple from inner query. Hence processing time will be $0(n^2)$ and $0(n) < 0(n)^2$ so processing time in (b) is higher than (a).

17. (b)
"Delete" is used to delete the data specified in the where clause. If where cause is absent then it will delete all the data from the table.
"Truncate table" is used to delete all the data but not able.
"Drop table" is used to delete the data as well as the table.

18. (d)
Given all statements are valid, as pairs given in each statement will return the same result.

19. (d)
All given query languages are equally expressive.

20. (d)
"Order by" clause is used to order the tuples and it can only be used as the last clause of a query.

23. (c)
The WHERE condition checks for the Grade = 'A' and matching of course-number in the courses and Grades, where the instructor must be 'Korth'. Overall it checks for students to whom 'Korth' has instructed and they have received Grade = 'A' in atleast one of the courses taught by 'Korth'.

24. (b)
Option (b) is correct as 'IN' condition will give the respective S.No. and will equate both S_No. But EXISTS will check if any tuple exists and it will give wrong answer.

25. (c)
For relation X cardinality = 2
For relation Y cardinality = 4
Hence the SQL statement
Select COUNT (*) from X, Y
gives the number of rows for the relation obtained from cross product of X, Y.
Hence, 2 × 4 = 8 is the correct option.

26. (b)
Query-2 will give the correct result but not query-1. Query-1 will check for employees who are working for department other than 5.
Query-2 will check for all employees even for department 5 as well.

27. (a)
Select R.did from case C, reserves R
where R.cid = C.cid and C.color = 'red' ...(1)
Query (1) will give:

did	22	31	64	22	31

Select R.did from case C, reserves R where R.cid = C.cid and C.color = 'green' ...(2)
Query (2) will give:

did	22	31	74

(1) intersect (2) will give:

did	22	31

The corresponding names would be : Karthikyan and Boris.

28. (b)

where competes.cid = competitor.cid

It gives

Cid
01
02
03
04

and nationality = 'Swedish'

It gives

Cid
03

So corresponding eid = 02.

29. (15)

This query counts the number of tuples with marks > minimum marks. All the tuples will have uniques marks. Hence only one tuple with minimum marks. Therefore 15 will be the result of the query (15 records will have marks greater than minimum marks).

30. (b)

Even through the R and S may have common attributes it would give output as cross-product unless we give explicit condition for tuple matching.

To make it equivalent to natural join we have to use WHERE condition to match the common attributes.

31. (c)

First query will successfully execute and after executing this statement, attribute Did of Employee table with ID = 1 becomes 'NULL' which is foreign key from Department's Did. Now, second statement will not execute, because primary key cannot be NULL.

32. (15020)

Nested Loop algorithm will involve $n_r \times b_s + b_r$ block transfers.

n_r = # records in relation r, b_r = # blocks in relation r

n_s = # records in relation s, b_s = # blocks in relation s

Either $T1$ can be R or $T2$.

If R is $T1$ then total number of block access is $4000 \times 20 + 50 = 80050$

If R is $T2$ then total number of block access is $300 \times 50 + 20 = 15020$

Better one is the second case, total number of block accesses = (15020).

33. (234516)

- Take all tables listed in 2 compute their cross product
- Select rows of cross product that satisfy the condition in 3
- Group the selected rows by attribute in 4
- Select the groups that satisfy the condition in 5
- Project out the selected group attribute in 1
- List the result by attribute in 6

34. (b)

After executing above command table Q will be empty and table P contains (a, NULL) and (b, NULL).

35. (c)

Query computes name of persons who have travelled in every city where SBI is located.

36. (6)

Since all names are distinct and all have gate score < 500 all rows will come in result.

Student Name	GATE Score
Amjad	300
Sachin	300
Mukesh	150
Nagender	100
Anand	200
Sushil	250

37. (a)

Inner query finds all budgets of Kanpur city.

Outer query compare every record of project if budget greater than equal to inner budget the reject it.

So finally results: Name of projects whoes budget is less than every project in Kanpur city.

■■■■

5 CHAPTER

Transaction

Q.1 Consider three data items D_1, D_2 and D_3 and the following execution schedule of transactions T_1, T_2 and T_3. In the diagram, $R(D)$ and $W(D)$ denote the action reading and writing the data item D respectively

Time	T_1	T_2	T_3
		$R(D_3)$;	
		$R(D_2)$;	
		$W(D_2)$;	
			$R(D_2)$;
			$W(D_3)$;
	$R(D_1)$;		
	$W(D_1)$;		
			$W(D_2)$;
			$W(D_3)$;
		$R(D_1)$;	
	$R(D_2)$;		
	$W(D_2)$;		
		$W(D_1)$;	

(a) The schedule is serializable as T_2, T_3, T_1
(b) The schedule is serializable as T_2, T_1, T_3
(c) The schedule is serializable as T_3, T_2, T_1
(d) The schedule is not serializable

Q.2 Which of the following statements is/are correct?

1. Every view serializable schedule is conflict serializable.
2. Every conflict serializable schedule is view serializable.
3. Some view serializable schedules are conflict serializable.
4. No view serializable schedule is conflict serializable.

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

Q.3 For the given schedule, which of the following is true?

T_1	T_2	T_3	T_4
	read(A)		read(A)
write(B)		read(A)	
	write(A)		
		read(B)	
	write(B)		

(a) The schedule cannot be serialized
(b) The schedule is equivalent to T_3, T_4, T_1, T_2
(c) The schedule is equivalent to T_1, T_4, T_3, T_2
(d) The schedule is equivalent to T_2, T_3, T_1, T_4

Q.4 The goal of concurrency control in database systems is to

(a) Allow only those concurrent executions of transactions that are equivalent to some serial executions of these transactions
(b) Allow only these transactions to execute concurrently that do not access any common relation
(c) Execute transactions serially
(d) Lock and unlock relations

Q.5 Consider the following situations:

1. Transaction T_1 is waiting for transactions T_2 and T_3.
2. Transaction T_3 is waiting for transtraction T_2.
3. Transaction T_2 is waiting for transaction T_4.

Above situation may result into

(a) Starvation of T_1
(b) Deadlock state
(c) No deadlock state
(d) Data not sufficient

Q.6 What is the potential problem when a DBMS executes multiple transactions concurrently?

(a) The lost update problem
(b) The dirty read problem

(c) The incorrect summary problem
(d) All of the above

Q.7 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

Q.8 **Assertion (A):** The following schedule is not conflict serializable schedule.
Reason (R): There is not any cycle in this schedule.

Time	T_1	T_2	T_3
			$R(D)$
		$R(A)$	
	$R(C)$		
		$W(A)$	
		$R(B)$	
		$R(C)$	
			$R(E)$
	$R(A)$		
			$R(B)$
		$R(D)$	
	$W(A)$		
			$W(B)$
	$R(E)$		

(a) Both (A) and (R) are true and (R) is the correct reason for (A)
(b) Both (A) and (R) are true but (R) is not correct reason for (A)
(c) Both (A) and (R) are false
(d) (A) is true but (R) is false

Q.9 Consider the following two schedules:

1.

T_1	T_2
$W_1(a)$	
	$R_2(b)$
$R_1(b)$	
	$R_2(a)$
C_1	
	C_2

2.

T_1	T_2	T_3
$R_1(a)$		
	$W_2(a)$	
	C_2	
$W_1(a)$		
C_1		
		$R_3(a)$
		C_3

Which of the following statement is true?
(a) 1 is recoverable, but 2 is not recoverable
(b) 1 is not recoverable but 2 is recoverable
(c) Both 1 and 2 are recoverable
(d) Both 1 and 2 are not recoverable

Q.10 Consider three data items D_1, D_2 and D_3 and the following execution schedule of transactions T_1, T_2 and T_3. In the diagram $R(D)$ and $W(D)$ denote the actions reading and writing the data item D respectively.

Time	T_1	T_2	T_3
		$R(D_3)$;	
		$R(D_2)$;	
		$W(D_2)$;	
			$R(D_2)$;
			$W(D_3)$;
	$R(D_1)$;		
	$W(D_1)$;		
			$W(D_2)$;
			$W(D_3)$;
		$R(D_1)$;	
	$R(D_2)$;		
	$W(D_2)$;		
		$W(D_1)$;	

(a) The schedule is serializable as T_2, T_3, T_1
(b) The schedule is serializable as T_2, T_1, T_3
(c) The schedule is serializable as T_3, T_2, T_1
(d) The schedule is not serializable

Q.11 A race condition occurs when
(a) Two concurrent activities interact to cause a processing error
(b) Two users of the DBMS are interacting with different files at the same time
(c) Two uses trying to obtain write lock on same data item
(d) None of these

Q.12 Consider the following schedules, which of the following schedule is conflict serializable?
(a) T_1 : $R(x)$, T_2 : $R(x)$, T_1 : $W(x)$, T_1 : commit, T_2 : commit
(b) T_1 : $R(x)$, T_2 : $R(y)$, T_3 : $W(x)$, T_2 : $R(x)$, T_1 : $R(y)$, T_1 : commit, T_2 : commit, T_3 : commit
(c) T_1 : $W(x)$, T_2 : $R(x)$, T_1 : $W(x)$, T_2 : commit, T_2 : Abort
(d) T_1 : $R(x)$, T_2 : $W(x)$, T_1 : $W(x)$, T_3 : $R(x)$, T_1 : commit, T_2 : commit, T_3 : commit

Q.13 Consider the following two transaction schedule:

Schedule-1

T_1	T_2	T_3	T_4
RA			
RB			
WA		RC	
	RA		RB
	WA		
WB			
		WC	
			WB
C			
	C		
		C	
			C

Schedule-2

T_1	T_2	T_3	T_4
RA			
			RC
	WA		
WB			
		WA	
			WB
WC			
C			
	C		
		C	
			C

Then which of the above schedule is serial equivalent?

(a) Both schedule 1 and 2
(b) Only Schedule 1
(c) Only Schedule 2
(d) None of these

Q.14 Consider the following schedule:

$R_1(X), R_1(Y), W_2(X), W_1(X), W_3(X);$

Which of the following is true regarding the given schedule?

(a) The schedule is conflict serializable but not view serializable
(b) The schedule is view serializable but not conflict serializable
(c) The schedule is view serializable as well as conflict serializable
(d) The schedule is neither view nor conflict serializable

Q.15 Assume there are 10 transaction : $T_1, T_2, \ldots T_{10}$. Every transaction writes a value of variable X that is read by its successive transaction, like T_1 writes a value of X that is read by T_1, T_2 writes a value of X, that is read by T_3, ..., T_9 writes a value of X that is read by T_{10}. Now, if T_1 fails, then total how many transactions are required to recover the schedule?

Q.16 For the given schedules S find out the conflict equivalent schedule:

S: $r_1(x)$; $r_2(z)$; $r_1(z)$; $r_3(x)$; $r_3(y)$; $w_1(x)$; $w_3(y)$; $r_2(y)$; $w_2(z)$; $w_2(y)$

(a) $T_3 \rightarrow T_1 \rightarrow T_2$ (b) $T_1 \rightarrow T_2 \rightarrow T_3$
(c) $T_2 \rightarrow T_1 \rightarrow T_3$ (b) None of these

Q.17 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

Q.18 Consider the following three schedules of transactions T1, T2 and T3.

[***Notation:*** In the following NYO represents the action Y (Y for read, W for write) performed by transaction N on object O]

(S1) 2RA 2WA 3RC 2WB 3WA 3WC 1RA 1RB 1WA 1WB
(S2) 3RC 2RA 2WA 2WB 3WA 1RA 1RB 1WA 1WB 3WC
(S3) 2RA 3RC 3WA 2WA 2WB 3WC 1RA 1RB 1WA 1WB

Which of the following statements is TRUE?

(a) S1, S2 and S3 are all conflict equivalent to each other
(b) No two of S1, S2 and S3 are conflict equivalent to each other
(c) S2 is conflict equivalent to S3, but not to S1
(d) S1 is conflict equivalent to S2, but not to S3

Q.19 When n transactions are run concurrently and in an interleaved manner, the number of possible schedules are _____.

(a) much larger than $n!$
(b) much lower than $n!$
(c) much larger than $(n-1)!$
(d) much lower than $(n-1)!$

Q.20 Which of the following is true?

(a) Every schedule which is not conflict serializable but view serializable must have some bliend write.

(b) Every schedule which is not conflict serializable but view serializable may/may not bliend write.

(c) Every schedule which is not conflict serializable and no bliend writes can be view serializable.

(d) Every schedule which is conflict serializable and no bliend writes is may not view serializable.

Q.21 Consider the following schedule.

$W_4(y), R_1(x), R_1(w), W_1(w), W_2(x), C_2, W_3(w), C_3, W_4(w), C_4, W_1(z), C_1$.

Which of the following transaction order is not possible in the given serializable schedule?

(a) $T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_4$

(b) $T_1 \rightarrow T_3 \rightarrow T_2 \rightarrow T_4$

(c) $T_1 \rightarrow T_3 \rightarrow T_4 \rightarrow T_2$

(d) $T_1 \rightarrow T_2 \rightarrow T_4 \rightarrow T_3$

Q.22 Consider the following schedule.

$R_1(x); R_2(x); W_1(x); R_2(y); W_1(x); C_2, A_1;$

The above schedule is

(a) Recoverable and conflict serializable

(b) Recoverable and not conflict serializable

(c) Not recoverable and conflict serializable

(d) Not recoverable and not conflict serializable

Q.23 Consider the given schedule

T_1	T_2
$R(x)$	
	$R(x)$
	$R(y)$
	$x = x + y$
	$W(x)$
$R(x)$	
	commit
commit	

The given schedule is

(a) Recoverable and cascadeless

(b) Recoverable but not cascadeless

(c) Not recoverable

(d) Recoverable, cascadeless and strict

Q.24 Find the schedule from the following which is serializable but not recoverable schedule? [C_1 is commit operation of transaction 1].

(a) $R_1(a), W_1(a), R_2(a), W_2(a), C_1, C_2$

(b) $R_1(a), R_2(a), W_2(a), W_1(a), C_1, C_2$

(c) $R_1(a), W_1(a), R_2(a), W_2(a), C_2, C_1$

(d) $R_1(a), W_2(a), W_1(a), R_2(a), C_2, C_1$

Q.25 Match List-I with List-II and select the correct answer using the codes given below the lists:

List-I

A. $W_1(x); R_2(x);$

B. $R_1(x); W_2(x);$

C. $W_1(x); W_2(x);$

List-II

1. Dirty read
2. Overwriting uncommited data
3. Unrepeatable read

Codes:

	A	B	C
(a)	1	2	3
(b)	1	3	2
(c)	2	1	3
(d)	2	3	1

Q.26 Match List-I with List-II and select the correct answer using the codes given below the lists:

List-I (Schedule name)

A. Recoverable

B. Csecade less

C. Strict

D. Non-recoverable

List-II (Schedule with example)

1. $R_1(A), W_2(B), R_1(B), C_1, C_2;$
2. $W_1(A), W_2(A), C_2, W_1(A), C_1;$
3. $R_1(A), W_2(B), C_2, R_1(B), W_1(B), C_1;$
4. $R_1(A), W_2(B), R_1(B), C_2, C_1;$

Codes:

	A	B	C	D
(a)	1	2	3	4
(b)	4	3	2	1
(c)	4	2	3	1
(d)	1	2	4	3

Q.27 Consider the following schedules:

S1: $W_2(x), W_1(x), R_3(x), R_1(x), W_2(y), R_3(y), R_3(z), R_2(x)$

S2: $R_3(z), W_2(x), W_2(y), R_1(x), R_3(x), R_2(z), R_3(y), W_1(x)$

S3: $R_2(z), W_2(x), W_2(y), W_1(x), R_1(x), R_3(x), R_3(z), R_3(y)$

Which of the schedules are serializable?

(a) S1, S2 (b) S1, S3
(c) S2, S3 (d) None of these

Q.28 Match **List-I** with **List-II** and select the correct answer using the codes given below the lists:

List-I

A. WR conflict
B. RW conflict
C. WW conflict

List-II

1. Unrepeatable read
2. Phantom read
3. Lost update
4. Reading uncommitted data

Codes:

	A	B	C
(a)	4	1	3
(b)	4	2	3
(c)	2	4	3
(d)	2	1	4

Q.29 Which of the following is true about the schedule S?

T_1	T_2	T_3
R(X)		
	R(X)	
R(Z)		
		R(X)
		R(Y)
W(X)		
Commit		
		W(Y)
		Commit
	R(Y)	
	W(Z)	
	W(Y)	
	Commit	

(a) Irrecoverable and cascading
(b) Recoverable and cascading
(c) Recoverable and cascadeless
(d) Irrecoverable and cascadeless

Q.30 Consider the following schedule S_1.

S_1: $R_1(C)\ R_2(C)\ W_1(A)\ W_2(A)\ W_1(C)\ R_1(B)\ R_2(B)\ W_1(B)\ W_1(D)\ W_2(B)\ W_2(D)\ R_1(F)\ W_3(E)\ R_3(F)$

Let 'X' be the number of 'blind-write' operations in the given schedule and 'Z' be the number of conflict equivalent serial schedules to S_1. The value of $X * Z$ is_____.

Q.31 Consider the transactions T_1, T_2, and T_3 and the schedules S_1 and S_2 given below.

T_1 : r1(X); r1(Z); w1(X); w1(Z)

T_2 : r2(Y); r2(Z); w2 Z)

T_3 : r3(Y); r3(X); w3(Y)

S_1 : r1(X); r3 (Y); r3(X); r2(Y); r2(Z); w3(Y); w2(Z); r1(Z); w1(X); w1(Z)

S_2 : r1(X); r3(Y); r2(Y); r3(X); w2(Z); r1(Z); r2(Z); w3(Y); w1(X); w1(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.

Q.32 Consider the following schedule

S: $r_2(X), r_1(Y), w_2(X), r_2(Y), r_3(X), w_1(Y), w_3(X), w_2(Y)$

How many minimum number of moves (where a move consists of changing the position of one of the operations) are required to convert S into a conflict serializable schedule?

Q.33 Consider following transactions

1. $r_1(A)\ w_1(A)\ r_1(B) \cdot w_1(B)$
2. $w_2(B) \cdot w_2(A)$

How many non-serial schedules between T_1 and T_2 which are serializable?

(a) 1 (b) 2
(c) 4 (d) 0

Q.34 Consider 3 transactions T_1, T_2 and T_3 having 2, 3 and 4 operations respectively. Find number of concurrent schedules.

Q.35 Consider a set of 2 transactions T_1 and T_2 with number of operations as 6 and 7 respectively. The number of concurrent schedules possible are

(a) 13! (b) 1716
(c) $6! \times 7!$ (d) None of these

Q.36 Consider a simple check pointing protocol and the following set of operations in the log.

(Start, T_4); (Start, T_1); (Write, T_1, a, 2, 3); (Write, T_4, b, 1, 2) (commit, T_1); (Start, T_3) (check point); (Write, T_3, C, 4, 7) (commit, T_4)

If crash happens now and the system tries to recover using both undo and redo operations, what are the contents of the undo list and redo list

(a) Undo: $T_1 T_3$; redo : T_4
(b) Undo: T_3; redo : T_4
(c) Undo: T_2; redo : T_2
(d) Undo: T_4; redo : T_1

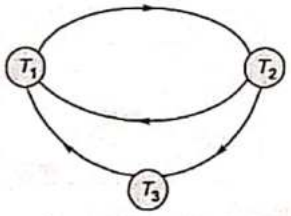
■■■■

Answers Transaction

1.	(d)	2.	(b)	3.	(c)	4.	(a)	5.	(c)	6.	(d)	7.	(d)	8.	(a)	9.	(c)
10.	(d)	11.	(a)	12.	(b)	13.	(d)	14.	(b)	16.	(a)	17.	(d)	18.	(d)	19.	(b)
20.	(a)	21.	(d)	22.	(a)	23.	(b)	24.	(c)	25.	(b)	26.	(c)	27.	(c)	28.	(a)
29.	(c)	31.	(c)	33.	(d)	35.	(b)	36.	(b)								

Explanations Transaction

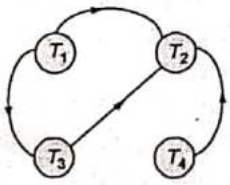
1. (d)
The precedence graph of the given schedule is



Since there is a cycle in the precedence graph, the schedule is not serializable.

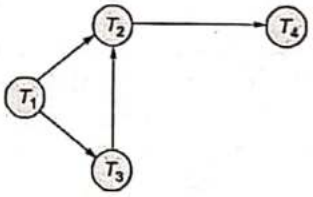
2. (b)
Every conflict serializable schedules are view serializable but vice-versa is not true. However some view serializable scheduls are conflict serializable.

3. (c)
The precedence graph of the given schedule is



Therefore schedule is equivalent to T_1, T_4, T_3, T_2 and T_1, T_3, T_4, T_2

5. (c)
Precedence graph for the situation is

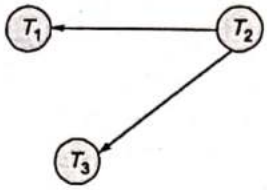


Since the graph has no cycle, the system is not in a deadlock state.

6. (d)
All the problems are potential when a DBMS executes multiple transactions concurrently.

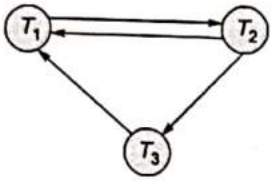
7. (d)
Irrecoverable error would be when a transaction reads a data item after it is written by a transaction which has not been committed yet.

8. (a)



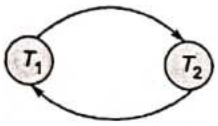
10. (d)

Because there exists a cycle in precedence graph.



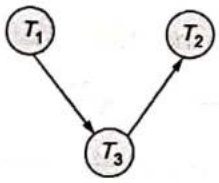
12. (b)

(a)



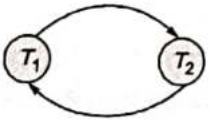
T_1	T_2
$R(x)$	
	$R(x)$
$W(x)$	
	$W(x)$

(b)



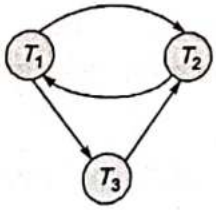
T_1	T_2	T_3
$R(x)$		
	$R(y)$	$W(x)$
	$R(x)$	
$R(y)$		

(c)



T_1	T_2
$W(x)$	
	$R(x)$
$W(x)$	

(d)

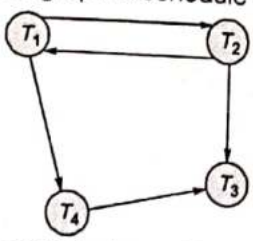


T_1	T_2	T_3
$R(x)$		
	$W(x)$	
$W(x)$		
		$R(x)$

If precedence graph consists cycle it is not conflict serializable otherwise it is conflict serializable.

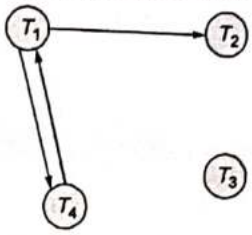
13. (d)

Precedence graph of schedule 1



There is a cycle between T_1 and T_2

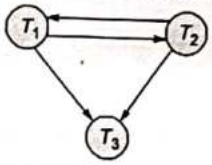
Procedure graph of schedule 2



There exists a cycle between T_1 and T_4.

14. (b)

There exists cycle in precedence graph



and

T_1	T_2	T_3
$R(X)$		
$R(Y)$		
	$W(X)$	
$W(X)$		
		$W(X)$

View Serializable schedule is

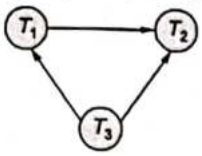
T_1	T_2	T_3
$R(X)$		
$R(Y)$		
$W(X)$		
	$W(X)$	
		$W(X)$

15. (10)

In order to recover the schedule, all the 10 transactions are required as the successive transaction is dependent on the data updated by previous transaction like T_2 is dependent on a variable which is updated by T_1. As T_1 fails, T_2 is also required to be recovered. Now if T_2 is recovered, T_3 is also recovered as T_3 depends on data updated by T_2. Likewise, it happens for all the transactions till T_{10}.

16. (a)

$S: r_1(x)\ r_1(z)\ r_3(x)\ r_3(y)\ w_1(x)\ w_3(y)\ r_2(y)\ w_2(z)\ w_2(y)$



The topological sorting / ordering of above graph is: $T_3 \rightarrow T_1 \rightarrow T_2$

This is the conflict equivalent schedule.

18. (d)

S_1

T_1	T_2	T_3
	R(A)	
	W(A)	
	W(B)	R(C)
		W(A)
		W(C)
R(A)		
R(B)		
W(A)		
W(B)		

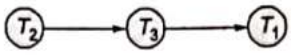
S_2

T_1	T_2	T_3
	R(A)	R(C)
	W(A)	
	W(B)	W(A)
R(A)		
R(B)		
W(A)		W(C)
W(B)		

S_3

T_1	T_2	T_3	
	R(A)		
		R(C)	
	W(A)	W(A)	Conflict pair
	W(B)		
		W(C)	
R(A)			
R(B)			
W(A)			
W(B)			

As we can see S_1 and S_2 are conflict equivalent to each other. The equivalent serial schedule is



While the schedule S_3 is having a conflicts pairs and equivalent serial schedule is not possible. Hence option (d) is correct.

19. (b)

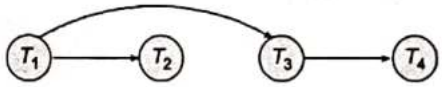
When n transactions are running concurrently and in an interleaved manner, the number of possible schedules are around $n!$.

20. (a)

If schedule is not conflict serializable but view serializable than their must be a blind write.

T_1	T_2
R(a)	W(a)
W(a)	Blind write

21. (d)



The possible orders:

$T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_4$

$T_1 \rightarrow T_3 \rightarrow T_4 \rightarrow T_2$

$T_1 \rightarrow T_3 \rightarrow T_2 \rightarrow T_4$

$\therefore\ T_1 \rightarrow T_2 \rightarrow T_4 \rightarrow T_3$ is not possible, $T_4 \rightarrow T_3$ is not serializable.

So option (d) is correct.

22. (a)

Schedule is recoverable, because there is no dirty read. If there exist a dirty read (*W-R*), then commit after parent of dirty read. i.e., if T_2 reads after T_1 writes over same data item x then T_2 should commit after T_1 commits.

$\Rightarrow$ Schedule is conflict serializable [not forms a cycle]

23. (b)

Recoverable but not cascadesless. For cascadeless T_1 should read the value of x only after commit operation of T_2.

Recoverable because T_2 commits first.

24. (c)

(a) Serializable [$T_1 \rightarrow T_2$]
Recoverable [T_2 reads 'a' after T_1 writes, hence T_2 commits after T_1]

(b) Not serializable [$T_1 \rightarrow T_2$, $T_2 \rightarrow T_1$]
Recoverable [No read after write]

(c) Serializable [$T_1 \rightarrow T_2$]
Not recoverable [T_2 commits before T_1]

(d) Not serializable and not recoverable

$\therefore$ Option (c) is serializable but not recoverable schedule.

25. (b)

Dirty read is W-R conflict [$A-1$]
Overwriting uncommited data is W-W conflict [$C-2$].
Unrepeatable read is R-W conflict [$B-3$].

26. (c)

(a) Recoverable:
$R_1(A), W_2(B), R_1(B), C_2, C_1$; [4]

(b) Casecadeless:
$W_1(A), W_2(A), C_2, W_1(A), C_1$; [2]

(c) Strict:
$R_1(A), W_2(B), C_2, R_1(B), W_1(B), C_1$; [3]

(d) Non recoverable:
$R_1(A), W_2(B), R_1(B), C_1, C_2$; [1]

Note: Strict schedule is recoverable and casecadeless schedule.
Casecadeless schedule is always recoverable.

27. (c)

S2: $T_1 \leftarrow T_2 \rightarrow T_3$: $T_2 \rightarrow T_3 \rightarrow T_1$

S3: $T_1 \leftarrow T_2 \rightarrow T_3$: $T_2 \rightarrow T_1 \rightarrow T_3$

$\therefore$ $S2$ and $S3$ are serializable. [$S1$ forms a cycle between T_1 and T_2]

28. (a)

The following conflicts leads to respective problems:

WR conflict : Reading uncommitted data.
RW conflict : Unrepeatable read.
WW conflict : Overwriting uncommitted data.

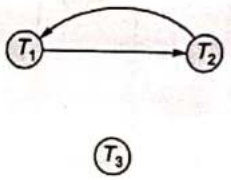
29. (c)

A schedule is said to be irrecoverable if T_i reads A which is updated by uncommitted T_j and commit of T_i is before T_j.
The given schedule does not have uncommitted read. So it is recoverable and cascadeless recoverable.

30. (0)

Blind-write: Write operation on a data item without having a read operation on the same data item before.
The following are 5 blind-writes:
$W_1(A), W_2(A), W_1(D), W_2(D), W_3(E)$
$\therefore$ $X = 5$

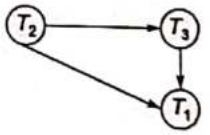
The precedency graph for above schedule S_1 is shown below:



Since cycle exist in the graph, the number of conflict serializable schedules are zero i.e. $Y = 0$.
$\therefore$ $X \cdot Y = 0$.

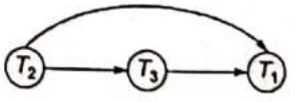
31. (c)

S_1: $r1(X); r3(Y); r3(X); r2(Y); r2(Z); w3(Y); w2(Z); r1(Z); w1(X); w1(Z)$



No cycle $\Rightarrow$ S1 is conflict serializable.

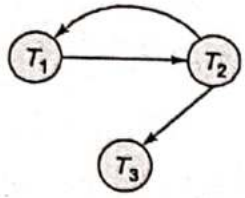
S_2: $r1(X); r3(Y); r2(Y); r3(X); w2(Z); r1(Z); r2(Z); w3(Y); w1(X); w1(Z)$



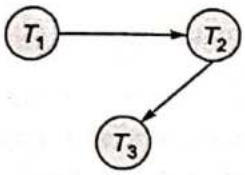
No cycle $\Rightarrow$ S_2 is conflict serializable.

32. (2)

Precedence Graph:



If $w_1(Y)$ is shifted before $r_2(Y)$. So first we have to move $w_1(Y)$ before $r_2(X)$, then $w_1(Y)$ before $r_2(Y)$ then graph becomes.



33. (d)

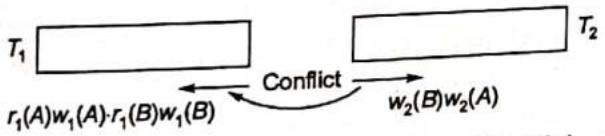
Number of concurrent schedule between T_1 and T_2

$$\frac{(4+2)!}{4!2!} = \frac{6!}{4!2!} = 15$$

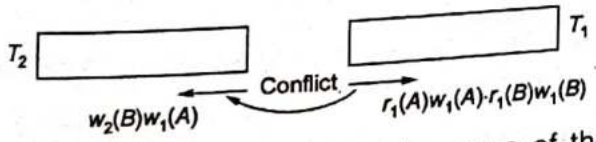
Number of serial schedule = 2!

$\because$ Number of non-serial schedule = 15 – 2 = 18

But, serializable schedule equal to $T_1 \rightarrow T_2$ serial



Serializable schedule equal to $T_2 \rightarrow T_1$ serial



Out of 13 non-serial schedule, none of the schedule equal to serial schedule.

$\because$ Number of non-serial schedule which are serializable = 0.

34. (1260)

If T_1, T_2 and T_3 are transactions with p, q and r operations respectively

Number of concurrent schedules = $\frac{(p+q+r)!}{p!\,q!\,r!}$

$$= \frac{(2+3+4)!}{2!\,3!\,4!} = \frac{9!}{2\times 6\times 4!}$$

$$= \frac{9\times 8\times 7\times 6\times 5}{2\times 6} = 63\times 20$$

$$= 1260$$

35. (b)

T_1 has 'n' operations and T_2 has 'm' operations, then number of concurrent schedules are

$$\frac{(n+m)!}{n!m!} = \frac{13!}{6!\times 7!}$$

$$= 13\times 11\times 3\times 4$$

$$= 1716$$

36. (b)

In data base until check point not come data is not saved permanently, when checkpoint is comes all database until checkpoint all data stored permanently.

After checkpoint process which are committed are redo and which are not committed are undo:

So, undo is to be transaction T_3

redo is to be transaction T_4

■■■■

6 CHAPTER
Concurrency Control Techniques

Q.1 The concept of locking can be used to solve the problem of
(a) Lost update
(b) Uncommited dependency
(c) Inconsistent data
(d) All of the above

Q.2 Assume transaction *A* holds a shared lock on *R*. If transaction *B* also requests for a shared lock on *R*, What will happen if rigrous two phase locking protocol is used?
(a) Results in a deadlock situation
(b) Immediately be granted
(c) Immediately be rejected
(d) Be granted as soon as it is released by *A*

Q.3 Which of the following is a false statement?
(a) Multiple granularity protocol is a variation of 2 Phase Locking protocol.
(b) Wait-Die protocol prevents dead lock and starvation.
(c) A schedule *S* is recoverable if no transaction *T* in *S* commits until all transaction *T'* which writes an item that *T* reads, have committed.
(d) None of the above.

Q.4 In wait-die scheme, tranasactions T_1 and T_2 have timestamps 7 and 9 respectively. If T_1 requests a data item held by T_2 then
(a) T_1 will wait
(b) T_1 will be rolled back
(c) T_2 will wait
(d) T_2 will be rolled back

Q.5 In wound-wait scheme, transactions T_1 and T_2 have timestamps 7 and 9 respectively. If T_1 requests a data item held by T_2 then
(a) T_1 will wait
(b) T_1 will be rolled back
(c) T_2 will wait
(d) T_2 will be rolled back

Q.6 Which of the following statements is true?
(a) Wait-die is a preemptive technique
(b) Wound-wait is a non-premptive technique
(c) Wait-die and would-wait are the schemes for deadlock recovery
(d) Wound-wait and wait-die schemes avoid stravation

Q.7 Consider the schedule given below. T_1 and T_2 are two transactions and x and y are two resources.

T_1	T_2
read_lock(y);	
read_item(y);	
	read_lock(x);
	read_item(x);
write_lock(x);	
	write_lock(y);

The schedule indicates
(a) serializable schedule
(b) non-serializable schedule
(c) A deadlock schedule
(d) None of the above

Q.8 Which of the following is/are true for transaction processing?
1. A transaction is a program unit where execution preserve the consistency of the database
2. Two-phase locking is used to prevent unauthorised user access to a database record
3. When a read-lock is acquired for a given record all other transactions will not have rights to read that record.
4. A record of all transactions and the corresponding changes to the database is recorded in a log

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

Q.9 Consider the following schedule:
$I_1(A)$, $R_1(A)$, $u_1(A)$, $I_2(A)$, $W_2(A)$, $u_2(A)$, $I_1(A)$, $W_1(A)$, $u_1(A)$.

Identify the schedule ?
(a) Schedule satisfies 2PL and conflict serializable
(b) Satisfies 2PL and non-conflict serializable
(c) Not satisfies 2PL and conflict serializable
(d) Not satisfies 2PL and not conflict serializable

Q.10 Match **List-I** (Simultaneously held locks on some object) with **List-II** (Group-mode for locks, if they are compatible) and select the correct answer using the codes given below the lists:

List-I	List-II
A. *S, S, IS*	1. *S*
B. *IX, IS, IS*	2. *IX*
C. *IX, X*	3. *SIX*
D. *SIX, IS*	4. Not possible

Codes:

	A	B	C	D
(a)	1	2	3	4
(b)	1	2	4	3
(c)	1	3	2	4
(d)	1	3	4	2

Q.11 Which of the following is true?
(a) Schedules which are allowed under Thomas write rule are also allowed under Basic times stamp protocol
(b) All the schedules which are allowed under basic time stamp protocol are also allowed under multiversion timestamp protocol
(c) All the schedules which are allowed under multi version time stamp protocol are also allowed under Thomas write rule
(d) All of these

Q.12 Consider the statement with regards to time-stamp ordering protocols
1. Starvation to a particular transaction is possible.
2. It ensures recoverable schedule.
3. It ensures conflict serializable schedule.
4. Will not allow dead locks.

Which of the above statements are true?
(a) 1 and 2 (b) 1, 2 and 4
(c) 1, 3 and 4 (d) All

Q.13 Which of the following is true for two-phase locking?
(a) lock acquisition is the second phase
(b) locks can be acquired at any time
(c) locks are acquired in the first phase
(d) None of the above

Q.14 Consider the following statements.
S_1 : Entire database can not be locked.
S_2 : Entire relation can be locked.

Which of the above statements is/are true?
(a) S_1 only
(b) S_2 only
(c) Both S_1 and S_2 are true
(d) both S_1 and S_2 are false

Q.15 Choose the false statement.
(a) Timestamp protocol is deadlock free.
(b) Two phase locking guarantees serializability.
(c) Strict two phase locking is deadlock free.
(d) Timestamp protocol may not result recoverable schedule.

Q.16 Which of the following concurrency control protocol ensures both conflict serializability and freedom from deadlock?
(i) 2 phase locking
(ii) Time stamp ordering
(a) (i) only
(b) (i) and (ii)
(c) (ii) only
(d) None of the above

Q.17 Suppose that process *P* has been running for several days when a new process *Q* starts up and begins contending with *P* for resources. Which of the following is true?
(a) In a wait-die system, if *P* needs a resource held by *Q*, then *P* waits.

(b) In a wait-die system, if Q needs a resource held by P, then Q waits.
(c) In wound-wait system, if P needs a resource held by Q, then Q yields and waits.
(d) In a wound-wait system, if Q needs a resource held by P, then Q dies.

Q.18 For the schedule given below, which of the following is correct?

1 Read A
2 Read B
3 Write A
4 Read A
5 Write A
6 Write B
7 Read B
8 Write B

(a) This schedule is serialisable and can occur in a scheme using 2PL protocol.
(b) This schedule is serialisable but cannot occur in a scheme using 2PL protocol.
(c) This schedule is not serialisable but can occur in a scheme using 2PL protocol.
(d) This schedule is not serialisable and cannot occur in a scheme using 2PL protocol.

Q.19 Which of the following is true about 2-phase locking protocol?

S_1: Lock upgradation and degradation are allowed only in shrinking phase.
S_2: 2-phase locking allows lock degradation in shrinking phase.

(a) Only S_1 (b) Only S_2
(c) Both S_1 and S_2 (d) Neither S_1 nor S_2

Q.20 2PL guarantees serializability, but it does not prevent deadlocks. 2PL has two phases : growing and shrinking. Which of the following rules are used to govern the 2PL protocol?

(a) Two transactions can not have conflicting locks.
(b) No unlock operation can precede a lock operation in the same transaction.
(c) No data are affected until all locks are obtained i.e, until the transaction is in its locked point.
(d) All of these

Q.21 Consider the following two transactions : T_1 and T_2.

T_1:	T_2:
$R(A)$;	$R(B)$;
$R(C)$;	$R(C)$;
$R(B)$;	$R(A)$;
If $A = 0$ then $B \leftarrow B \times 5$;	If $B \neq 0$ then $A \leftarrow A / 5$;
$W(B)$;	$W(A)$;

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

(a)

S1 :	S2 :
$S(A)$;	$S(B)$;
$R(A)$;	$R(B)$;
$S(C)$;	$S(C)$;
$R(C)$;	$R(C)$;
$S(B)$;	$S(A)$;
$R(B)$;	$R(A)$;
if $A = 0$	if $B \neq 0$
then $B \leftarrow B \times 5$;	then $A \leftarrow A / 5$;
$W(B)$;	$W(A)$;
C;	C;
unlock (A);	unlock (B);
unlock (C);	unlock (C);
unlock (B);	unlock (A);

(b)

S1 :	S2 :
$X(A)$;	$X(B)$;
$R(A)$;	$R(B)$;
$S(C)$;	$S(C)$;
$R(C)$;	$R(C)$;
$X(B)$;	$X(A)$;
$R(B)$;	$R(A)$;
if $A = 0$	if $B \neq 0$
then $B \leftarrow B \times 5$;	then $A \leftarrow A / 5$;
$W(B)$;	$W(A)$;
unlock (A);	unlock (A);
unlock (C);	unlock (C);
C;	C;
unlock (B);	unlock (A);

(c)

S1 :	S2 :
$S(A)$;	$S(B)$;
$R(A)$;	$R(B)$;
$S(C)$;	$S(C)$;
$R(C)$;	$R(C)$;

$X(B)$;	$X(A)$;
$R(B)$;	$R(A)$;
if $A = 0$	if $B \neq 0$
then $B \leftarrow B \times 5$;	then $A \leftarrow A / 5$;
$W(B)$;	$W(A)$;
unlock (A);	unlock (A);
unlock (C);	unlock (C);
unlock (B);	unlock (B);
C;	C;

(d)

$S1$:	$S2$:
$S(A)$;	$S(B)$;
$R(A)$;	$R(B)$;
$S(C)$;	$S(C)$;
$R(C)$;	$R(C)$;
$X(B)$;	$X(A)$;
$R(B)$;	$R(A)$;
if $A = 0$	if $B \neq 0$
then $B \leftarrow B \times 5$;	then $A \leftarrow A / 5$;
$W(B)$;	$W(A)$;
unlock (A);	unlock (B);
unlock (C);	unlock (C);
C;	C;
unlock (B);	unlock (A);

Q.22 Consider the following sequence of actions.

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

Which of the following time stamp ordering allowed to execute the schedule using basic time stamp protocol?

(a) $(T_1, T_2, T_3) = (10, 30, 20)$
(b) $(T_1, T_2, T_3) = (20, 30, 10)$
(c) $(T_1, T_2, T_3) = (30, 20, 10)$
(d) $(T_1, T_2, T_3) = (10, 20, 30)$

Q.23 Consider the following schedule:

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

Which of the following time stamp ordering allows to execute schedule using Thomas write rule time stamp ordering protocol?

(a) $(T_1, T_2, T_3) = (10, 30, 20)$
(b) $(T_1, T_2, T_3) = (20, 10, 30)$
(c) $(T_1, T_2, T_3) = (30, 20, 10)$
(d) $(T_1, T_2, T_3) = (10, 20, 30)$

■■■■

Answers Concurrency Control Techniques

1.	(d)	2.	(d)	3.	(d)	4.	(a)	5.	(d)	6.	(c)	7.	(c)	8.	(c)	9.	(d)
10.	(b)	11.	(b)	12.	(c)	13.	(c)	14.	(b)	15.	(a)	16.	(c)	17.	(a)	18.	(b)
19.	(b)	20.	(b)	21.	(d)	22.	(a)	23.	(d)								

Explanations Concurrency Control Techniques

4. (a)
In wait-die scheme, 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. Thus when T_1 requests a data item held by T_2, then T_1 will wait.

5. (d)
In wound-wait scheme, when transaction T_i requests a data item currently held by T_j, then T_i is rolled back if timestamp of T_j is larger than T_i. Thus for the given question, data item will be preempted from T_2 and T_2 will be rolled back.

6. (c)
Wait -die is a non-preemptive technique. Wound-wait is a preemptive technique. Wait-die and Wound-wait are the schemes for deadlock prevention. Both the schemes avoid starvation.

7. (c)
Transaction T_1 is waiting for data item x, which is held by T_2. And Transaction T_2 is waiting for data item y, which is held by T_1. Both transactions need Exclusive locks so the situation is in deadlock state.

9. (d)

After unlock over A, it locks again on A.

Hence it is not in 2PL.

$\therefore$ not conflict serializable.

10. (b)

- S, S, IS can held with shared [1]
- IX, IS, IS can held with IX [2]
- IX and X are incompatible, so not possible [4]
- SIX, IS can held with SIX [3]

16. (c)

Time stamp ordering ensure both conflict serializability and free from deadlock.

19. (b)

According to 2PL, if lock conversion is allowed, then upgrading of locks must be done during expansion phase, and degrading of locks must be done in the shrinking phase.

$\therefore$ S_2 is true.

Lock upgradation is not allowed in shrinking phase therefore S_1 is False.

20. (b)

No unlock operation can precede a lock operation in the same transaction.

21. (d)

Requirement to follow Strict 2PL:

1. Exclusive locks should be released after the commit.
2. No Locking can be done after the first Unlock and vice versa.

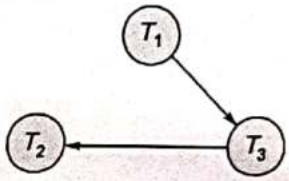
In 2PL, deadlock may occur but it may be that it doesn't occur at all.

Consider that in option (d) if both execute in serial order without concurrency then that is perfectly valid and YES it follows Strict 2PL.

22. (a)

Time stamp order is equal to the conflict equal serial order.

Precedence graph:



23. (d)

T_1 (10)	T_2 (20)	T_3 (30)
$r(A)$		
	$r(B)$	
$w(C)$		$w(B)$
		$r(C)$
	$w(B)$	
		$w(A)$

$W_3(B)$, $W_2(B)$ is allowed in *TWR*.

■■■■

7 CHAPTER
File Organization and Indexing

Q.1 Match **List-I** with **List-II** and select the correct answer using the codes given below the lists:

List-I	List-II
A. Primary index	1. Dense
B. Clustered index	2. Sparse
C. Secondary key index	
D. Secondary non-key index	

Codes:

	A	B	C	D
(a)	2	1	1	2
(b)	1	1	2	1
(c)	1	2	2	1
(d)	2	2	1	1

Q.2 Blocking factor is applicable for _____ organization

(a) Spanned (b) Un-spanned
(c) Both (d) None of these

Q.3 Each block of *B* tree of order '*n*' can have

(a) $n-1$ key values, and n pointers
(b) $n+1$ key values, and n pointers
(c) n key values, and $n+1$ pointers
(d) Cannot be determined

Q.4 Dense index is so called because

(a) It normally keep one key per data block
(b) Every key from the data file is represented in the index
(c) It requires many fewer blocks than the sparse index
(d) None of these

Q.5 Duplicates are allowed in

(a) *B* Trees (b) B^+ Trees
(c) Both (a) and (b) (d) None of these

Q.6 Which statement is true for a B tree?

(a) All entries of a node are greater than or equal to the entries in the node's children.
(b) All leaves are at the exact same depth.
(c) All nodes contains the exact same no of entries.
(d) All non-leaf nodes have the exact same number of children.

Q.7 The physical block is the smallest physically addressable unit of data on a disk. The disk block addresses are used to access a physical record. The blocking address consists of

(a) Cylinder number (b) Surface number
(c) Block number (d) All of the above

Q.8 B tree allows

(a) Sequential access only
(b) Random access only
(c) Both sequential and random access
(d) None of these

Q.9 **Assertion (A):** Second level index in multilevel indexing will be always a primary index.
Reason (R): All indexed files will be always physically ordered files.

Which one of the following is true?

(a) Both (A) and (R) are true and (R) is the reason for (A)
(b) Both (A) and (R) are true but (R) is not correct reason for (A)
(c) Both (A) and (R) are false
(d) (A) is true but (R) is false

Q.10 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 point is 6 bytes. The largest possible order of a non-leaf node is a B^+ tree implementing this file structure is

(a) 23 (b) 24
(c) 34 (d) 44

Q.11 Consider a table T in a relational database with a key field K. A B 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 index node. Assume that k is 10 bytes long; disk block size is 512 bytes; each data pointer P_D is 8 bytes along 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

(a) 20 (b) 22
(c) 23 (d) 32

Q.12 Given a data file with N (e.g. 100) records per page and M (e.g. 1000) pages and an index page capacity of 512 index entries, how deep should be the B^+ tree to index this file?

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

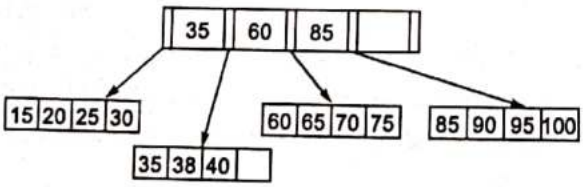
Q.13 Match **List-I** with **List-II** and select the correct answer using the codes given below the lists:

List-I	List-II
A. Secondary index dependency	1. Functional
B. Non-procedural Query language	2. B Tree
C. Closure of a set of attributes	3. Domain calculus
D. Natural-join algebraic operations	4. Relational

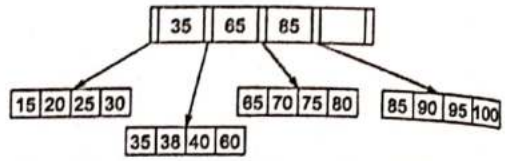
Codes:

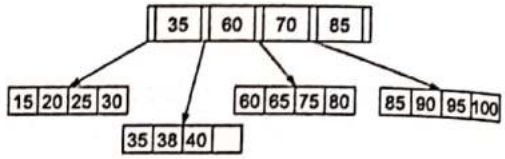
	A	B	C	D
(a)	2	3	1	4
(b)	3	2	4	1
(c)	3	1	2	4
(d)	2	1	4	3

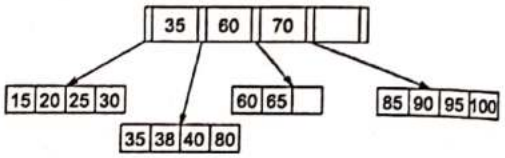
Q.14 Consider the following B+ tree



What will be the resulting B^+ tree after inserting 80. Assume B^+ tree incorporate rotation technique

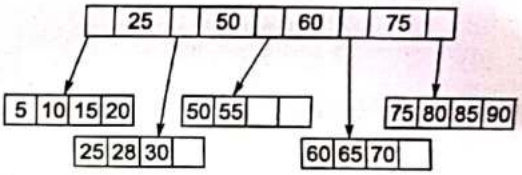
(a)


(b)


(c)


(d) None of the above

Q.15 Consider the following B^+ tree with the corresponding guidelines



Guidelines

- Number of keys/page = 4
- Number of pointers/page = 5
- Fill factor = 50%
- Minimum keys in each page = 2
- While combining a leaf page, combine it with its left sibling
- Rotation technique of inserting is not allowed
- The root node and Intermediate nodes are index pages.
- In case when leaf page is full, then while splitting the leaf page middle key is placed in the index page in sorted order. Records with keys < middle key go to left leaf page. Records with keys ≥ middle key go to the right leaf page.
  While index page is full then while splitting index page, keys < middle key go to left index page and keys > middle key go to the right index page. The middle key goes to the next (higher level) index. If key value 95 is added to the above B^+ tree the root page of the resulting B^+ tree will contain

(a) 60 (b) 25, 50, 60, 75
(c) 25, 50, 75, 80 (d) 80

Q.16 Suppose that we store records on a disk device having the following characteristics:
Average access motion time : 0.02 sec.
Disk rotation speed: 3,000 revolutions per minute
Data transfer rate: 250,000 bytes per sec

What is the expected data transfer time (in sec.) for a randomly accessed physical record that is 500 bytes in length?

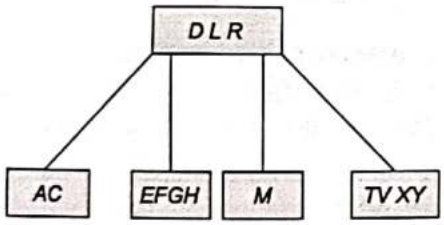
(a) 0.0212 (b) 0.0230
(c) 0.0240 (d) 0.0320

Q.17 Which of the following file organization provides very fast access to any arbitary record of a file?

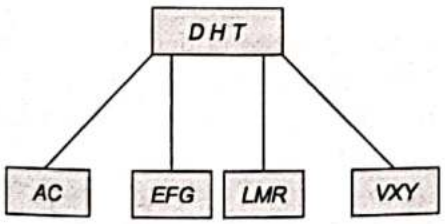
(a) Ordered file (b) Unordered file
(c) Hashed file (d) B^+ tree

Q.18 Identify the *B*-tree from the following after inserting *R, Y, F, X, A, M, C, D, E, T, H, V, L, G* (in the order) with branching factor of 3.
If any insertion of clement causes split, first split that node without element and after balanced then insert that element.

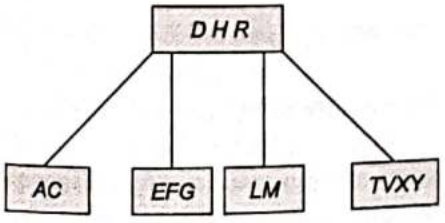
(a)



(b)



(c)

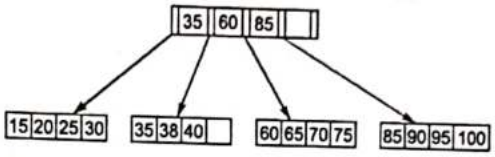


(d) None of these

Q.19 In a *B*-Tree, the block size is 512 bytes, search key is 2 bytes, block pointer size is 6 bytes and record pointer size is 8 bytes. Every node contains $n-1$ record pointers, $n-1$ search keys, n block pointers. Find out the average number of search keys that can be accommodated if the fill factor of B-tree is 67%?

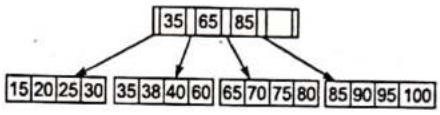
(a) 20 (b) 21
(c) 31 (d) None of these

Q.20 Consider the following B^+ tree:

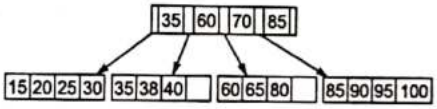


What will be the resulting B^+ tree after inserting 80. Assume B^+ tree incorporate rotation technique.

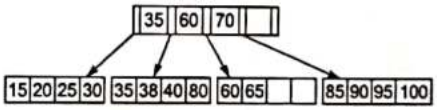
(a)



(b)



(c)



(d) None of these

Q.21 Consider the following data:
A file has 2^{29} records each of size 8 B. One block of main memory is 128B. Sparse indexing is done with one index record per memory block and one index record is of 1 byte. The blocks stored by the index records are stored in disk. Blocks occupied by index are searched using binary search technique. If on the disk system block read takes 10 m sec, then binary search will take

(a) 110 m sec (b) 120 m sec
(c) 150 m sec (d) 180 m sec

Q.22 Assume a system has the following parameters to construct a B^+ tree:

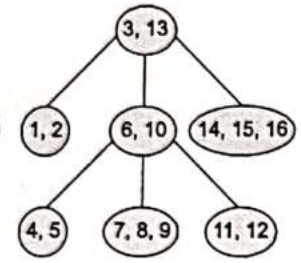
- System block size of 232 bytes,

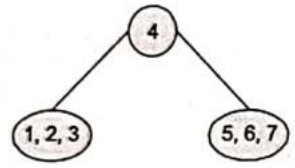
- Block pointers are of 12 bytes each
- Record pointers are of 12 bytes each
- Index keys are of 8 bytes each

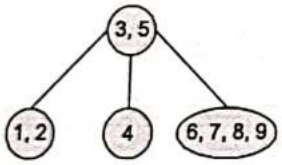
Find the maximum order of the *B*+ tree, if the block can contain at most one *B*+ node in the system?

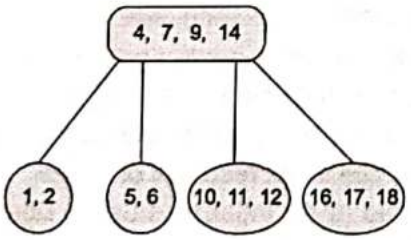
(a) 9 (b) 10
(c) 11 (d) 12

Q.23 Which of the following is a legal B-tree ? Assume minimum branching factor is 3.

(a)


(b)


(c)


(d)


Q.24 What is the "maximum height" of *B*-tree, if a *b*-tree has *n*-keys with degree *t*.

Note: Node can contain more than one key other than root.

(a) $\left\lfloor \log_t^{n/2} \right\rfloor$ (b) $\left\lfloor \log_t^{n+1} \right\rfloor$
(c) $\left\lfloor \log_t^{n} \right\rfloor$ (d) $\left\lfloor \log_1^{(n+1)/2} \right\rfloor$

Q.25 Consider the following structure of an student record:

```
struct student
{   int sno;
    char sname[22];
    float marks;
    char subj[10];
};
```

Assume integer, character and float data of 2 bytes, 1 byte and 4 bytes respectively. Block size of 512 bytes is used in unspanned organization and each file has 500 records. Find the unused space in the last block of a file?

(a) 18 bytes (b) 156 bytes
(c) 266 bytes (d) 284 bytes

Q.26 What is primary distinction between *B*-tree indices and *B*+ tree indices?

(a) *B*-tree eliminates the redundant storage of search-key values
(b) *B*+ - tree eliminates the redundant storage of search key indices
(c) Deletion in a *B*-tree is more complicated
(d) Lookup in *B*-tree is proportional to the logarithm of the number of search key

Q.27 The record pointer, key field and block pointer of a '*B*' tree are 8B, 10B and 6B respectively. Calculate the order of the tree if the block size is 1 kB

(a) 44 (b) 43
(c) 24 (d) None of these

Q.28 Consider *B* tree of degree '*n*' with height '*h*'. Calculate maximum number of records that can be indexed by the above tree (assume root is at level '0')

(a) $n^h - 1$ (b) n^h
(c) $n^{h+1} - 1$ (d) None of these

Q.29 Suppose we have a block addressable disk drive with such block organized disk non data overhead of sub-blocks and inter-block gaps have to be accounted for, there are 40,000 bytes per track and the amount of space taken up by sub-blocks and inter-blocks 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 orgnaised to hold an integral number of logical records.

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

Q.30 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.31 Which of the following is not a file operation?

(a) FindNext (b) Find
(c) Reset (d) None of these

Q.32 Index sequential file is made of all of these expect

(a) primary data storage area
(b) overflow area
(c) hierarchy of indices
(d) address of prime data track

Q.33 Which allocation scheme would work best for a file system implemented on a device that can only be accessed sequentially, a tape drive, for instance?

(a) Contiguous allocation
(b) Non Contiguous allocation
(c) Indexed allocation
(d) None of the above

Q.34 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.35 Which indices search key defines the sequential order of file and which indices search key specifies an order different from sequential order of file?

(a) Primary, Secondary
(b) Clustering, Primary
(c) Non-clustering, Secondary
(d) Clustering, Secondary

Q.36 Which of the following is correct?

(a) B tress 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.

Q.37 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.

(a) 50 (b) 51
(c) 1024 (d) 15

Q.38 A B tree used as an index for a large databse table has four levels inculding the root node. If a new key is inserted in this index, then the maximum number of nodes that could be newly created in the process are

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

Q.39 Given a block can hold either 3 records or 10 key pointers. A database contains n records, then how many blocks do we need to hold the data file and the dense index.

(a) $\frac{13n}{30}$ (b) $\frac{n}{3}$
(c) $\frac{n}{10}$ (d) $\frac{n}{30}$

Q.40 Which is the best suitable for sequential access of data

(a) B tree (b) B^+ tree
(c) both (d) None of these

Q.41 Consider unspanned blocking with 20 byte blocks. A file contains records of sizes 2, 5, 3, 7, 4, 20 bytes.
The percentage of space wasted if blocks are allocated for file is _____.

Q.42 Assume that search key size is 90% of the record size within a file (records are fixed length records). Pointer makes up only 10% of the record size. Based on the above assumptions which of the following statements is true?

(a) Index block factor >> Database file block factor
(b) Database file block factor >> Index block factor
(c) Database file block factor = (Index block factor)/2
(d) Database file block factor = Index block factor

Q.43 The maximum number of records that can be indexed for a B^+ tree of 4 levels with order 10 and root at level 1 are _____.

Q.44 A file is organized so that the ordering of data record is same as or close to the ordering of data block in some index. Then that index is called

(a) Dense (b) Sparse
(c) Clustered (d) Unclustered

Q.45 Consider a *B*-tree of order 15 with 3 levels [Root is at level 1]. The maximum possible no of keys which can be indexed in the above B tree is _____.

Q.46 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 3 bytes, the search field value takes 17 bytes, and the block size is 1024 bytes. The order of the internal node is _____.

Q.47 The following key values are inserted into a B^+ - tree in which order of the internal nodes is 4, and that of the leaf nodes is 3, 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.

3, 10, 12, 14, 29, 38, 45, 55, 60, 68

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

Q.48 Consider a Table *T* with a key field K. *A B* tree of order P where P denotes the maximum number of record pointers in a *B* tree. Assume *K* is 12 bytes long, disk block size is 1024 Bytes, record pointer is 6 bytes and block pointer size is 2 bytes long. In order for each *B* tree node to fit in a single disk block the maximum value of *P*.

Q.49 Consider the following statement:

(*i*) Primary index is always sparse
(*ii*) Secondary index may or may not be dense

Which of the above statement is/are false?

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

Q.50 The following key values are inserted into an empty B^+ tree where each node have two key values. Find sum of keys present at height one (assume root is present at height 0)

8, 5, 1, 7, 3, 12, 9, 6

Q.51 A block can hold either 12 records or 24 key pointers. A database contains 96 records, then how many blocks are required to hold the data file and the dense index?

(a) 10 (b) 12
(c) 11 (d) 13

Q.52 Suppose a phone book contain 500 pages and each page can contain upto 500 records. Suppose we want to search for a particular name in a phone book. Give a worst case bound on number of pages that must be looked to perform a search using an index for the name of the first entry of each page.

Q.53 A data base table T_1 has 1000 records and occupies 40 disk blocks. Another table T_2 has 200 records and occupies 10 disk blocks. These two tables have to be joined as per specified join condition that needs to be evaluated for every pair of records from these two tables. The available memory space can accommodate one block of each relations. No index available for any table.

The maximum number of block access required for reading the data if nested loop join is used are _____.

Q.54 Consider a relation R with 6 attributes $A_1, A_2, A_3, A_4, A_5, A_6$. Given the following FD's

$$A_1 \rightarrow A_2$$
$$A_2 A_3 \rightarrow A_5$$
$$A_5 A_6 \rightarrow A_4 A_1$$

The number of candidate keys that includes A_2 as attributes _____.

Q.55 In a database file, the search field is 8 bytes long the block size is 512 bytes. 8 record pointer is 7 bytes and block pointer is 6 bytes. The largest possible order of a non leaf node in B^+ tree implementing this file structure?

■■■■

Answers File Organization and Indexing

1.	(d)	2.	(b)	3.	(a)	4.	(b)	5.	(b)	6.	(b)	7.	(d)	8.	(b)	9.	(c)
10.	(c)	11.	(c)	12.	(b)	13.	(a)	14.	(a)	15.	(a)	16.	(d)	17.	(c)	18.	(c)
19.	(a)	20.	(a)	21.	(d)	22.	(c)	23.	(b)	24.	(d)	25.	(d)	26.	(d)	27.	(b)
28.	(c)	29.	(c)	30.	(c)	31.	(a)	32.	(d)	33.	(d)	34.	(a)	35.	(d)	36.	(d)
37.	(b)	38.	(b)	39.	(a)	40.	(d)	42.	(d)	44.	(b)	49.	(c)	51.	(c)		

Explanations File Organization and Indexing

10. (c)

$p \times$ size of tree pointer + $(p - 1) \times$ (size of key field) $\leq$ Block size

Since it is B^+ tree. In order for each B^+ tree node to fit in a single disk block the maximum value of p is

$$p \times P + (p - 1) V \leq B$$
$$p \times 6 + (p - 1) \times 9 \leq 512$$
$$15p \leq 521$$
$$p \leq 34$$

11. (c)

Typical structure of a B-tree non-leaf node is follows:

P_1	B_1	K_1	P_2	B_2	K_2	–	P_{m-1}	B_{m-1}	K_{m-1}	P_m

in order for each B-tree node to fit in a single disk block the maximum value of p is.

$(p \times P) + ((p - 1) \times (p_r + V)) \leq B$

where disk block size B = 512 byte

block pointer p = 5 B

record (data pointer) P = 8 B

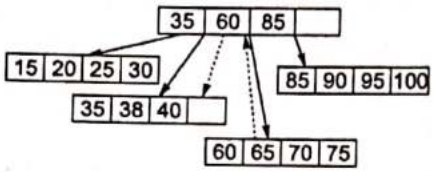
search field K or V = 10B

$$(p \times 5) + ((p - 1) \times (8 + 10)) \leq 512$$
$$5p + 18p - 18 \leq 512$$
$$23p \leq 512 + 18$$
$$p \leq \left[\frac{530}{23}\right] = 23$$

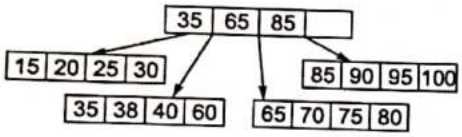
12. (b)

One level B^+ tree node will have 512 index entries which could only refer to 512 pages. A two level B^+ tree will have 512^2 index entries at the leaf level. There are 262,144 entries. This is enough to address the 1000 pages of the file. Moreover, sufficient to address each record individually, since there are only 100,000 records.

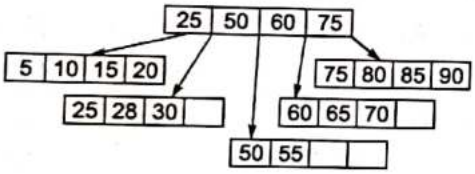
14. (a)



B^+ tree incorporate rotation technique.
Insertion of 80 would be done between 60 and 85. But in the node between 60 and 85, there is no space and also the child node of 85 and above, has no space left. So due to rotation technique, 60 will be inserted into the child node of 35 and 60. 65 will be shifted upwards and thus 80 will be inserted.
So the tree will look like;



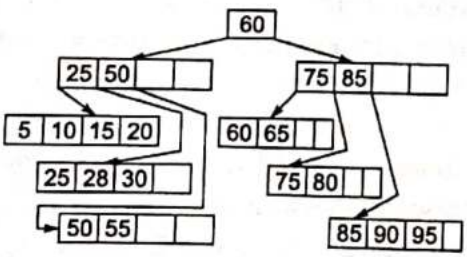
15. (a)



'95' will be inserted in node 75 80 85 90, but the node is full.
According to the properties given in que, the middle element of the node will be placed in next (higher) index while splitting.
So, in 75 80 85 90 95, 85 will be placed upwards.
Now, in above index, 25 50 60 75 the node is full, so again need to split up.

25 50 60 75 85

60 is middle element, so it will go to upper layer/ root index. So tree will be:



16. (d)
Average Access Time = Seek Time + Average Rotational Delay + Data Transfer Time
Seek Time = 0.02 (given)

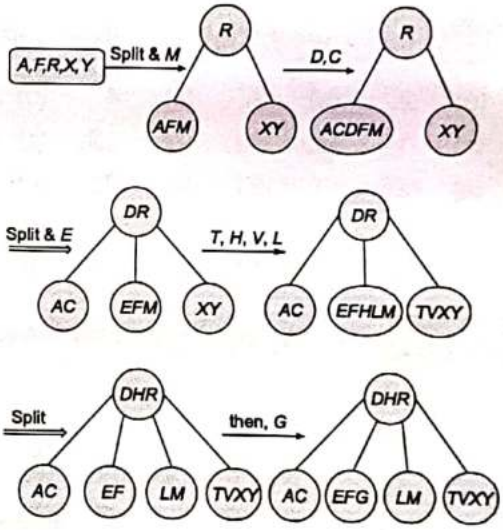
$$\text{Average Rotational Delay} = \frac{60}{3000} \times \frac{1}{2} = 0.01$$

$$\text{Data Transfer Time} = \frac{500}{250000} = 0.002$$

So Access Time = 0.02 + 0.01 + 0.02 = 0.320

17. (c)
Hashed file can be accessed with O(1) complexity.

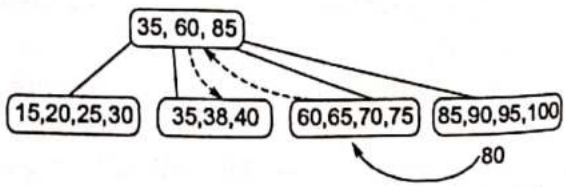
18. (c)



19. (a)

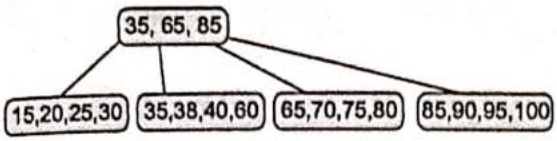
If fill factor is 67% so $32 \times \frac{67}{100} = 21.44 = 21$
So 21 block pointers
Number of keys = Number of block pointers – 1
i.e. (21 – 1) = 20 keys

20. (a)



Sorted order 60, 65, 70, 75, 80 as left leaf has space, so we can do redistribution.

So



21. (d)

$\because$ one block read takes 10 msec.

$\therefore$ 18 blocks read takes $18 \times 10 = 180$ msec.

22. (c)

keys (Index keysize + block pointer size) + Record pointer size $\le$ Blocksize

$\Rightarrow \quad n(8+12) + 12 \le 232$

$\Rightarrow \quad n \le \dfrac{232-12}{20}$

$\Rightarrow \quad n \le 11$

$\therefore$ Maximum order = 11

23. (b)

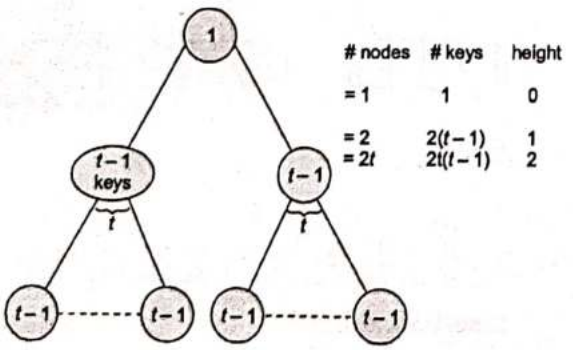
(a) All leafs are not at same level (not balanced). It is not *B*-tree.

(b) Root can have 1 key and remaining contains minimum 2 keys. Hence it is *B*-tree.

(c) "4" key has less than allowable size. (Minimum 2 keys)

(d) Right to 7, there is no key.

24. (d)



Max height is possible if every node has minimum keys.

Root can contain minimum 1 key

Other nodes can contain minimum $(t-1)$ keys

$\therefore \; n \ge 1 + 2(t-1) + 2t(t-1) + 2t^2(t-1) + \ldots + 2t^h(t-1)$

$$\ge 1 + 2(t-1)\sum_{x=1}^{h} t^{x-1}$$

$$\ge 1 + 2(t-1)\left(\frac{t^h-1}{t-1}\right)$$

$$n \ge 1 + 2(t^h - 1)$$

$\Rightarrow \quad 2t^h - 2 \le n - 1$

$\Rightarrow \quad t^h \le \dfrac{n+1}{2}$

$\Rightarrow \quad h \le \log_t^{(n+1)/2}$

$\Rightarrow \quad h = \lfloor \log_t^{(n+1)/2} \rfloor$

25. (d)

Number of blocks required for a file

$$= \left\lceil \frac{500 \text{ records}}{13 \text{ records/block}} \right\rceil = 39$$

39 blocks required for 500 records and each block occupied by 13 records.

$39 \times 13 - 500 = 7$ records space is not used for file. The last block holds only 6 records for file and remaining space is wasted.

$\therefore \; 512 - 6 \times 38 = 284$ bytes of space unused.

27. (b)

$nBp + (n-1)$ key $+ (n-1)$ record pointer $\le 1024B$

$n \times 6 + (n-1) \times 10 + (n-1) \times 8 \le 1024$

$6n + 10n + 8n - 18 \le 1024$

$24n \le 1024 + 18$

$$n \le \frac{1042}{24}$$

$n \le 43.41 \Rightarrow n = 43$

28. (c)

At level '0', the number of records $= n - 1$

At level '1', the number of records $= n(n-1)$

At level 'h' the number of records $= n^h(n-1)$

Total $= (n-1)(1 + n + \ldots n^{h-1} + n^h)$

$$= (n-1)\frac{(n^{h+1}-1)}{(n-1)} = n^{h+1} - 1$$

29. (c)

Number of blocks per track = 32

Number of bytes per track = 40000 bytes

$\therefore$ Number of bytes per block $= \dfrac{40000}{32}$

= 1250 bytes

Space taken by sub-blocks and inter blocks gap = 250 bytes

So remaining 1250 – 250 = 1000 bytes are used for records.

$\therefore$ record size = 200 bytes

$\therefore$ Number of records per block = $\frac{1000}{100} = 5$

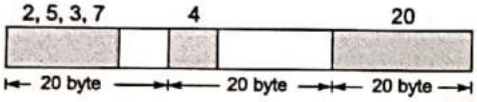
39. (a)

For storing the records number of blocks required $= \frac{n}{3}$ and for storing the keys in dense index

number of blocks required = $\frac{n}{10}$

So total blocks required = $\frac{n}{3} + \frac{n}{10} = \frac{13n}{30}$

41. (31.66)



Space wasted = 3 + 16 = 19

% of space wasted = $\frac{19}{60} \times 100 = 31.66\%$

42. (d)

It is given that key size + Record pointer size = Record size

This makes $\left\lfloor \frac{\text{Blocks size} - \text{Header}}{\text{Key} + \text{Record pointer}} \right\rfloor$

= Index block factor = $\left\lfloor \frac{\text{Blocks size} - \text{Header}}{\text{Record size}} \right\rfloor$

= Database file block factor

43. (9000)

In level 2 there will be 10 nodes as order is 10.
In level 3 there will be 10 × 10 = 100 nodes
In level 4 there will be 10 × 100 = 1000 nodes
At level 4 each node will have 9 record pointers.
So, therefore the maximum number of records that be indexed are 1000 × 9 = 9000.

44. (b)

- In sparse index number of entries depends on number of data block in previous level.
- In dense index number of entries depends on number of data record in previous level.

45. (3374)

Level 0 ↓ Level 1 ↓ Level 2 ↓

$14 + 15 \times 14 + 15 \times 15 \times 14 = 3374$

46. (52)

Size of child pointer = 3 byte
Size of search field value takes 17 bytes
Block size = 1024
The order of internal node = P
(∵ Number of block pointers in any node)

$\Rightarrow (P-1)17 + P*3 \le 1024$

$17P + 3P - 17 \le 1024$

$20P \le 1024 + 17$

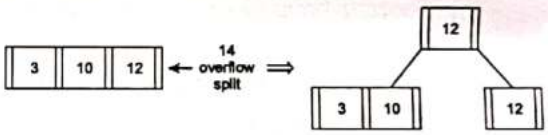
$P \le 1041/20$

$= \lfloor 52.05 \rfloor \approx 52$

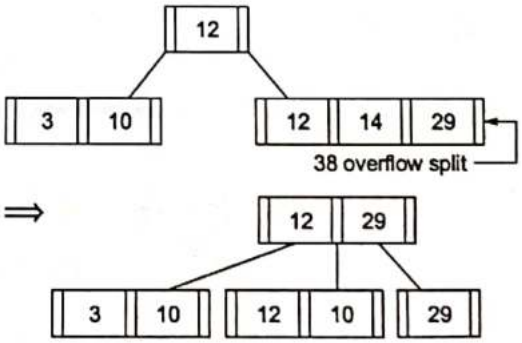
47. (4)

No. of key in internal node = 3 i.e, [4 – 1 = 3]
No. of key in leaf node = 3 i.e., [order = # Key]

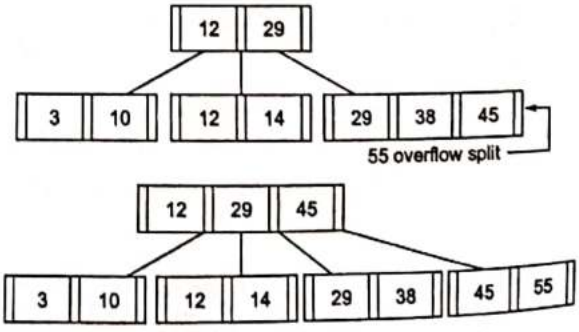
Insert 3, 10, 12:



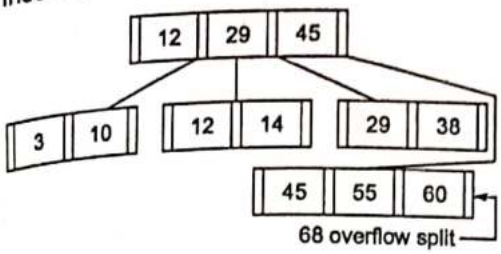
Insert 14, 29:



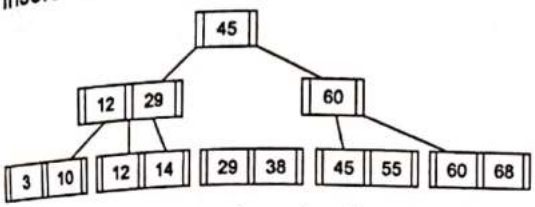
Insert 38, 45:



Insert 60:



Insert 68:



So the maximum number of splits of leaf node is 4.

48. (51)

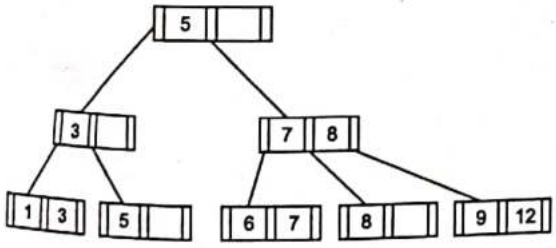
$$(12+6)P+(P+1)\cdot 2 \le 1024$$
$$18P+2P+2 \le 1024$$
$$20P \le 1022$$
$$P \le 51.1$$
$$P = 51$$

49. (c)

Primary index may or may not be sparse.

Secondary index is always dense.

50. (18)



51. (c)

To store the records, the number of blocks needed $= \frac{96}{12} = 8$.

To store the keys in dense index, the number of blocks needed $= \frac{96}{24} = 4$.

$\therefore$ Total number of blocks required $= 8 + 4 = 12$

52. (2)

If the book were organized as a heap file in a linear search so worst case bound is 500. By using binary search the worst case bound is $\lceil \log_2 500 \rceil = 9$ using the fact that the book is ordered by names with an index for the name of the first entry on each page we get worst case bound is 2 because the index has 500 pages, so entire index files in one page 1 access to the index and 1 to the data.

53. (10040)

In nested loop join for each records in table 1 all records of table 2 are loaded

So, two choice

1. T_1 as outer table
   $(1000 \times 10) + 40 = 10040$
2. T_2 as outer table
   $(200 \times 40) + 20 = 8020$

54. (1)

$$(A_2A_3)^+ \rightarrow A_3A_2A_5$$

So add A_6

$$(A_2A_3A_6)^+ \rightarrow A_2A_3A_5A_6A_1 \quad ...(1)$$
$$(A_1A_3)^+ \rightarrow A_1A_3A_2A_5$$

So add A_6

$$(A_1A_3A_6)^+ \rightarrow A_1A_2A_3A_5A_6A_4 \quad ...(2)$$
$$(A_3A_5)^+ \rightarrow A_3A_5$$

add A_6

$$(A_3A_5A_6)^+ \rightarrow A_3A_5A_6A_4A_1A_2 \quad ...(3)$$

Three candidate keys are $(A_2A_3A_6)$ $(A_1A_3A_6)$ and $(A_3A_5A_6)$ in which one candidate key contain A_2 as atrribute.

55. (34)

Order of B+ tree for non leaf node = Number of block pointers

$(P-1) * \text{Key} + P * \text{Block pointer} \le 512$

$(P-1) * 8 + P * 7 \le 512$

$15P \le 520$

$$P = \left\lfloor \frac{520}{15} \right\rfloor = 34$$

■■■■