



NPTEL Online Certification Courses

Indian Institute of Technology Kharagpur
Jan 2025



Course Name: Database Management System

Assignment 0 (Jan 2025)

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 20

Total mark: $20 \times 1 = 20$

Question 1

Look at the following truth table:

A	B	A op B
F	F	T
F	T	T
T	F	T
T	T	F

Which binary operation has been carried out?

- a) $\neg A \vee B$
- b) $\neg (A \vee B)$
- c) $\neg (A \wedge B)$ $\neg(A \wedge B) \equiv (\neg A) \vee (\neg B)$
- d) $\neg A \wedge B$

Answer: (c)

Explanation: The last column in the truth table shows that the result of the operation is true (or false) when A and B is false (or true).

Question 2

Consider the following array of seven integers:

30, 15, 10, 40, 25, 20, 12

What will be the contents of this array after the 2nd pass of bubble sort (sorting from smallest to largest)?

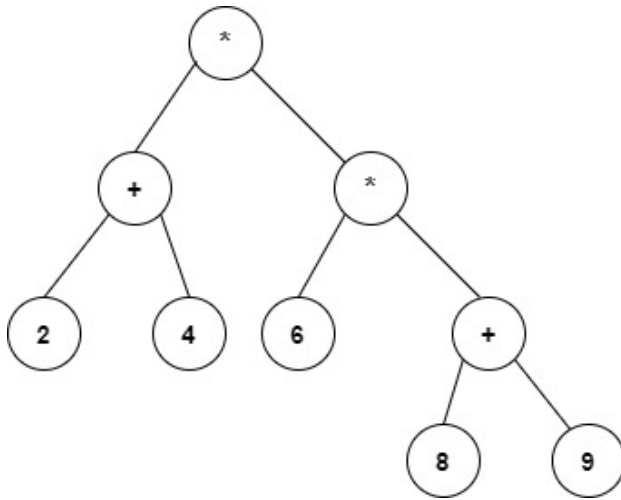
- a) 15, 10, 30, 25, 20, 12, 40
- b) 10, 12, 15, 25, 20, 30, 40
- c) 10, 15, 25, 20, 12, 30, 40
- d) 10, 12, 15, 20, 25, 30, 40

Answer: (c)

Explanation: According to the bubble sort algorithm.

Question 3

Identify the correct representation of the prefix expression of the given tree below.



- a) $*2 + 4 * 6 + 8 9$
- b) $2 + 4 * 6 * 8 + 9$
- c) $2 4 + 6 8 9 + **$
- d) $* + 2 4 * 6 + 8 9$

Answer:(d)

Explanation: Prefix expression is the pre order traversal fo the tree.

Question 4

Consider a set $X = \{ \{a, c\}, \{b, d\} \}$. What is the number of elements in the power set of X ?

- a) 4 The size of the power set of a set with n elements is given by:
- b) 8 $P(X) = 2^n$
- c) 12 where n is the number of elements in the set X
- d) 16

Answer: (a)

Explanation: Number of elements in X is 2. Hence, the number of elements in its power set is 2^2 . So, option (a) is correct.

Question 5

When an EVEN decimal number is converted into a binary number, what will be the the Least Significant Bit(LSB)?

- a) 0
- b) 1
- c) either 0 or 1
- d) only 11

Answer: (a)

Explanation: If any even decimal number is divided by 2, the remainder is 0 . This remainder will be the LSB of the binary equivalent. So, LSB of binary representation of any even decimal number will be always 0.

So, option a) is correct.

Question 6

Consider the following arrays of numbers. Which one represents a min-heap?

a)

20	5	15	10	6	8	9
----	---	----	----	---	---	---

b)

20	10	15	5	6	8	9
----	----	----	---	---	---	---

c)

9	6	10	8	5	15	20
---	---	----	---	---	----	----

d)

5	6	10	8	9	15	20
---	---	----	---	---	----	----

Complete Binary Tree – A min-heap is always a complete binary tree, meaning all levels are fully filled except possibly the last level, which is filled from left to right.

Heap Property – The value of each parent node is less than or equal to the value of its child nodes.

Root Contains Minimum Value – The smallest element is always located at the root node.

Answer: (d)

Explanation: According to the definition of the min-heap, any parent in a complete binary tree, must be lesser than both its children.

Hence, option (d) is correct.

A complete binary tree is a type of binary tree in which:

All levels are completely filled except possibly the last level.

In the last level, all nodes are filled from left to right without any gaps.

Question 7

Consider two sets A and B. Which of the following statement is true?

- a) $A - B = A \cap B$
- b) $A - B = A - (A \cap B)$
- c) $A - B = A - \neg B$
- d) $A - B = A \cap (A \cap B)$

Answer: (b)

Explanation: If we draw Venn Diagram, we shall see that the statement in option (b) is only correct.

Question 8

Let a given function be $f: \mathbb{I} \rightarrow \mathbb{I}$ given by $f(x) = x^2$ where \mathbb{I} stands for the set of all integers. Which of the following statements is/are true?

- a) It is a one-one function but not onto.
- b) It is an onto function but not one-one.
- c) It is both one-one and onto.
- d) It is neither one-one nor onto.

Answer: (d)

Explanation: By the given function, $f(-1)=1$ and $f(1)=1$. Hence it is not one-one. $f(x)$ should always be a positive integer but. $f: \mathbb{I} \rightarrow \mathbb{I}$. Consider a negative integer in the co-domain. There is no value of x for which $f(x)$ can be a negative integer.

Question 9

Consider the following:

$CSE(x)$: The girl x is in CSE department.

$DBMS(x)$: The girl x studies DBMS.

Consider the statement :

Which of the following formula represents “Not all CSE students study DBMS.”

a) $\forall(x) (CSE(x) \wedge \neg DBMS(x))$

b) $\forall(x) (\neg CSE(x) \rightarrow DBMS(x))$

c) $\exists(x) (\neg CSE(x) \rightarrow DBMS(x))$

d) $\exists(x) (CSE(x) \wedge \neg DBMS(x))$

Answer: (d)

Explanation: The first option is $\forall(x) (CSE(x) \wedge \neg DBMS(x))$. This is equivalent to $\forall(x) \neg (CSE(x) \rightarrow DBMS(x))$. That means CSE girls do not study DBMS.

The second option states that if the girl is not in CSE, she studies DBMS.

The third option states that some not CSE girls studies DBMS.

The fourth option states that not all CSE girls study DBMS.

Hence, option (d) is correct.

Question 10

Consider the binary relation $S_1 = \{(1, 1), (2, 2), (1, 2), (2, 1)\}$ on the set $\{1, 2\}$. Identify the correct statement.

- a) S_1 is reflexive.
- b) S_1 is symmetric.
- c) S_1 is not transitive.
- d) S_1 is antisymmetric.

Answer: a), b)

Explanation: S_1 reflexive, symmetric and transitive.

S_1 is not antisymmetric as both $(1, 2)$ and $(2, 1)$ are present in S_1

Question 11

What is the domain of the following function if the range is the set of real numbers?

$$\frac{5}{\sqrt{x^2 - 9}}$$

- a) $(-\infty, -3) \cup (-3, \infty)$
- b) $(-\infty, 3) \cup (3, \infty)$
- c) $(3, \infty) \cup (-\infty, -3)$
- d) $(-3, 3)$

Answer: c)

Explanation: The denominator is positive if x is not between -3 and +3. Hence, (c) is correct.

Question 12

How many functions are there from the set $\{1, 2, 3\}$ to the set $\{a, b\}$?

- a) 9
- b) 8
- c) 6
- d) 5

Answer: b)

Explanation: The element "1" can be mapped to any of the two elements in the range. Similarly, the element "2" or "3" can be mapped to any of the two elements in the range. Hence, (b) is the correct option.

Question 13

If U is the set of integers excluding zero, V is the set of even integers, and D is the set of odd integers, which of the following statement(s) is (are) true. Note: Here \bar{E} denotes complement over the expression E . U is the Universal set

a) $V \cup D = U$

b) $U - V = D$

c) $\bar{V} = D$

d) $V - D = \{0\}$

Answer: b), c)

Explanation: Follow the definition of universal set and difference operation.

Question 14

Suppose that A and B are two sets. Which of the following is always equivalent to $A \cup B$

a) $(A - B) \cup (B - A) \cup (A \cap B)$

b) $(A - B) \cup (B - A)$

c) $A \cup (B - A)$

d) $\bar{A} \cap \bar{B}$

Answer: a), c)

Explanation: As per the properties set operations, options (a) and (c) are correct.

Question 15

Consider the following function from $\{a,b,c,d\}$ to $\{1,2,3,4,5\}$.

$\{(a,2), (b,1), (c,3), (d,4)\}$

Which of the following is true about the function?

- a) The function is bijective
- b) The function is injective, but not surjective
- c) The function is surjective, but not injective
- d) The function is neither injective nor surjective

Answer: b)

Explanation: The function is one to one but not onto. Hence, option (b) is correct.

Question 16

Numbers 1, 2, 3, 4 are pushed into a stack in that order but these four PUSH operations are intermixed with POP operations as well. Whenever a number is popped, it is printed. Which of the following permutation can be printed by such PUSH and POP operations?

- a) 3, 4, 1, 2
- b) 1, 4, 2, 3
- c) 4, 2, 3, 1
- d) 3, 2, 4, 1

Answer: d)

Explanation: If 3 is printed first, 1, 2 are in the stack. Next, 4 is pushed and popped. But 1 cannot be popped now. Hence (a) is not possible.

1 is pushed and popped. Then, 4 is printed. So, 2, 3 are in the stack. 2 cannot be popped now. Hence (b) is not possible.

If 4 is printed first, 1, 2, 3 are in the stack. 2 cannot be popped now. Hence (c) is not possible.

If 3 is printed first, 1, 2 are in the stack. Next, 2 is popped. Then, 4 is pushed and popped. Lastly, 1 is popped. Hence (d) is the correct option.

Question 17

Which one of the following is the most appropriate logical formula to represent the statement?

“Red and Yellow cars are beautiful”.

The following notations are used:

$R(x)$: x is a Red car

$Y(x)$: x is a Yellow car

$B(x)$: x is beautiful

Note: $\forall x$ denotes for all x . $\exists x$ denotes for some x (atleast 1). \rightarrow denotes implication. \vee and \wedge denotes OR and AND operation respectively.

a) $\exists x(R(x) \vee Y(x)) \rightarrow B(x)$

b) $\exists x(R(x) \wedge Y(x)) \rightarrow B(x)$

c) $\forall x(R(x) \vee Y(x)) \rightarrow B(x)$

d) $\forall x(R(x) \wedge Y(x)) \rightarrow B(x)$

Answer: c)

Explanation: This statement can be expressed as

For all x , if x is either a red car or yellow car then the car x is beautiful.

Hence, option (c) is correct.

Question 18

Consider a B+-tree of order 5. What is the minimum number of children a root node can have, considering it is not the only node in the tree?

- a) 5
- b) 6
- c) 2
- d) 1

Answer: c)

Explanation: As per the construction rule of B+ tree, option (c) is correct.

Question 19

A hash table with length of 10 elements use open addressing with hash function $h(k)=k \bmod 10$, and linear probing. the table shows the state after insertion of 6 values. Identify a possible order in which the key values could have been inserted in the table.

0	
1	
2	22
3	43
4	14
5	62
6	96
7	103
8	
9	

- a) 96, 22, 14, 62, 43, 103
- b) 14, 22, 43, 62, 103, 96
- c) 22, 96, 43, 103, 14, 62
- d) 96, 14, 22, 43, 62, 103

Answer: d)

Explanation: In option a), 96 is placed at 6th index in the hash table. 22 is placed at 2nd index. 14 is placed at 4th index. When 62 is going to be placed in 2nd index, there is a collision. Therefore, 62 is placed at 3rd index. Hence, option (a) is not correct.

In a similar way, we can check for other options as well. We shall see that option (d) is correct.

Question 20

How many comparisons does it take to find the element 10 in the sorted array $A = \{10, 15, 17, 19, 20\}$ using binary search?

- a) 2
- b) 3
- c) 5
- d) 6

Answer: a)

Explanation: Using binary search technique:

First comparison with $\text{mid} = 17$

Second comparison with $\text{mid} = 10$ (this is the key).

Hence, option (a) is correct.



NPTEL Online Certification Courses

Indian Institute of Technology Kharagpur
Jan 2025



Course Name: Database Management System

Assignment 1 - Week 1 (Jan 2025)

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10

Total mark: $10 \times 2 = 20$

Question 1

Which of the following statements is (are) incorrect?

- a) Logical level abstraction defines the physical storage of data.
- b) View level abstraction provides a user-friendly interface for end users.
- c) Physical level abstraction focuses on the relationships among data.
- d) Logical level abstraction focuses on the relationships among data.

Answer: a), c)

Explanation: **Logical level:** describes the data stored in the database and the relationships among them.

Physical level: focuses on how data is stored physically.

View level: provides abstraction for users, hiding the complexities of the database structure.

Question 2

Consider the following SQL statement(s):

S1:

```
INSERT INTO employees(emp_id, emp_name, dept)
VALUES (101, 'Alice', 'HR');
```

S2:

```
ALTER TABLE employees ADD COLUMN salary number(8, 2);
```

Identify the correct statement.

- a) Both **S1** and **S2** are Data Manipulation (DML) Queries
- b) **S1** is a Data Manipulation (DML) Query, and **S2** is a Data Definition (DDL) Query
- c) Both **S1** and **S2** are Data Definition (DDL) Queries
- d) **S1** is a Data Control Query, and **S2** is a Data Definition (DDL) Query

Answer: b)

Explanation: Based on the syntax of the language, option (b) is the correct answer. DDL(Data Definition Language) includes CREATE, DROP, ALTER, TRUNCATE, RENAME; DML(Data Manipulation Language) includes INSERT, UPDATE, DELETE. For more details, refer to Module 2.

Question 3

Identify the valid **primary key** for the relation `event_registration` from the given instance.

event_registration			
participant_id	event_id	registration_date	status
1001	E001	2024-01-10	Confirmed
1002	E002	2024-01-11	Pending
1003	E001	2024-01-10	Confirmed
1001	E003	2024-01-12	Confirmed
1004	E002	2024-01-14	Pending

- a) `participant_id`
- b) `event_id, registration_date`
- c) `participant_id, event_id`
- d) `event_id, status`

Answer: c)

Explanation: A composite primary key consists of two or more attributes that together uniquely identify a record in the table. Option c) is the correct answer because the combination of `participant_id` and `event_id` ensures uniqueness for each registration. Other options do not guarantee uniqueness across all rows.

Feature	CHAR	VARCHAR
Length Type	Fixed-length	Variable-length
Storage Allocation	Allocates full size even if not used (pads with spaces)	Allocates only the space needed for the actual string
Performance	Faster for fixed-length data	Slightly slower due to dynamic length handling
Best Use Case	When string values have consistent lengths (e.g., country codes, fixed-size IDs)	When string values have varying lengths (e.g., names, descriptions)
Example	<code>CHAR(5)</code> → <code>'E123 '</code>	<code>VARCHAR(5)</code> → <code>'E123'</code>
Trailing Spaces Handling	Trailing spaces are retained	Trailing spaces are removed

Question 4

Identify the correct statement/s.

- a) `Employee(empID, empName)` is an instance of a relation schema.
- b) `Employee(empID, empName)` is an example of a physical schema.
- c) `(101, John)` is an instance of a relation schema.
- d) `(101, John)` is an example of a logical schema.

Answer: c)

Explanation: `(101, John)` is an instance of the schema `Employee(empID, empName)`. A relation schema defines the structure of the data, and an instance contains actual data values fitting the schema.

Concept	Definition	Example
Relation Schema	Structure of a relation (table)	<code>Employee(empID, empName)</code>
Instance of a Relation	Actual row in the table	<code>(101, John)</code>
Physical Schema	Storage details of the data	Use of B-tree or hash index
Logical Schema	Structure and constraints of the data	Table and attribute definitions

Question 5

Consider a relation `CityDetails(CityName, Population, CountryName)` where the superkeys are as follows: $\{CityName\}$, $\{CityName, Population\}$, $\{CityName, CountryName\}$, $\{CityName, Population, CountryName\}$.

Select the possible candidate key(s).

- a) $\{CityName, Population, CountryName\}$
- b) $\{CityName, Population\}$
- c) $\{CityName\}$
- d) $\{CityName, CountryName\}$

Answer: c)

Explanation: Minimal superkeys are candidate keys. Here, `CityName` alone is a superkey, so any superset of `CityName` will also be a superkey. But only `CityName` can be a candidate key. Hence, (c) is the correct option.

empID	empName	empAge	empDepartment
101	John	25	HR
102	Alice	30	IT
103	Bob	28	Finance

✓ Super Keys:

- `{empID}` → Unique
- `{empID, empName}` → Unique but not minimal
- `{empID, empAge}` → Unique but not minimal

✓ Candidate Keys:

- `{empID}` → Minimal and unique
- `{empName}` → If names are unique

Question 6

Consider the following relations:

Employee(eid, ename, salary)

Department(did, eid)

Consider the following Relational Algebras:

$RA_1 : \Pi_{\text{Employee.eid, ename}}(\text{Employee} \bowtie \text{Department})$

$RA_2 : \Pi_{\text{Employee.eid, ename}}(\text{Employee} \times \text{Department})$

Which of the following is correct?

a) $RA_1 = RA_2$

b) $RA_1 \subset RA_2$

c) $RA_2 \subseteq RA_1$

d) $RA_1 \subseteq RA_2$

Answer: d)

Explanation: A natural join only includes tuples where the **eid** matches, whereas a Cartesian product includes all possible combinations of tuples. Hence, option (d) is correct.

Question 7

Consider the following instance of the `CourseDetails(CourseID, CourseName)` relation:

CourseID	CourseName
C101	Mathematics
C102	Physics

If `CourseID` is the foreign key in the relational schema `StudentEnrollments(EnrollmentID, CourseID, StudentName)`, which of the following is a valid instance of `StudentEnrollments`?

a)

EnrollmentID	CourseID	StudentName
E001	C101	Amit
E002	C105	Raj

b)

EnrollmentID	CourseID	StudentName
E001	C101	Amit
E001	C102	Raj

c)

EnrollmentID	CourseID	StudentName
NULL	C102	Amit
E003	C102	Raj

d)

EnrollmentID	CourseID	StudentName
E001	C101	Amit
E002	C102	Raj

Answer: d)

Explanation: Option (a) is incorrect because `CourseID` “C105” does not exist in the referenced relation `CourseDetails`. Option (b) is incorrect because `EnrollmentID` cannot have duplicate values, violating the primary key constraint. Option (c) is incorrect because primary key values cannot be NULL, as seen in the first tuple. Hence, option (d) is the correct answer.

Question 8

Consider the following table:

CityDetails		
CityName	Population	StateName
Mumbai	20000	Maharashtra
Delhi	19000	Delhi
Bengaluru	12000	Karnataka
Hyderabad	10000	Telangana
Ahmedabad	8000	Gujarat
Pune	6000	Maharashtra

Identify the correct operation(s) that produce the following output from the above **relation**:

CityName	Population	StateName
Mumbai	20000	Maharashtra
Delhi	19000	Delhi
Bengaluru	12000	Karnataka

- a) $\sigma(\text{Population} \geq 12000)$ (CityDetails)
- b) $\sigma(\text{Population} > 10000) \wedge (\text{StateName} = \text{'Maharashtra'})$ (CityDetails)
- c) $\sigma(\text{Population} \geq 12000) \vee (\text{StateName} = \text{'Delhi'})$ (CityDetails)
- d) $\sigma(\text{Population} \geq 19000)$ (CityDetails)

Answer: a), c)

Explanation: As per Relational Operators syntax and semantics, options (a) and (c) correctly filters cities with a population greater than or equal to 12000 as per the given instance.

Question 9

Consider the following tables:

CityDetails ₁		
CityName	Population	StateName
Mumbai	20000	Maharashtra
Delhi	19000	Delhi
Ahmedabad	8000	Gujarat
Pune	6000	Maharashtra

CityDetails ₂		
CityName	Population	StateName
Mumbai	20000	Maharashtra
Bengaluru	12000	Karnataka
Ahmedabad	8000	Gujarat
Chennai	10000	Tamil Nadu

Identify the correct operation(s) which produce the following output from the above two relations:

CityName	Population	StateName
Mumbai	20000	Maharashtra
Ahmedabad	8000	Gujarat

- a) $\text{CityDetails}_1 - \text{CityDetails}_2$
- b) $\text{CityDetails}_1 \cap \text{CityDetails}_2$
- c) $\text{CityDetails}_2 - \text{CityDetails}_1$
- d) $\text{CityDetails}_1 \cup \text{CityDetails}_2$

For union there should be same number of columns in both tables

Answer: b)

Explanation: As per Relational Operators syntax and semantics, option (b) is correct because the result is the intersection of CityDetails_1 and CityDetails_2 .

Question 10

Consider the following table:

CityDetails		
CityName	Population	StateName
Mumbai	20000	Maharashtra
Delhi	19000	Delhi
Bengaluru	12000	Karnataka
Hyderabad	10000	Telangana
Ahmedabad	8000	Gujarat
Pune	6000	Maharashtra

Identify the correct operation(s) that produce the following output from the above **relation**:

FilteredCityDetails	
CityName	StateName
Mumbai	Maharashtra
Delhi	Delhi
Bengaluru	Karnataka

- a) $\sigma(\text{CityName}, \text{StateName})(\text{CityDetails})$
- b) $\sigma(\text{Population} \geq 12000)(\text{CityDetails})$
- c) $\Pi(\text{CityName}, \text{StateName})(\sigma(\text{Population} \geq 12000)(\text{CityDetails}))$
- d) $\Pi(\text{CityName}, \text{StateName})(\text{CityDetails})$

Answer: c)

Explanation: The operation in option (c) first applies the **selection operator** ($\sigma(\text{Population} \geq 12000)$) to filter rows where **Population** is at least 12000. Then, the **projection operator** ($\Pi(\text{CityName}, \text{StateName})$) is applied to retain only the **CityName** and **StateName** columns, which matches the desired output.



Course Name: Database Management System

Assignment 2 - Week 2 (Jan 2025)

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10

Total mark: $10 \times 2 = 20$

Question 1

In an e-commerce system, the instance of `OrderDetails` is as follows:

OrderDetails		
OrderID	ProductID	Quantity
0101	P001	3
0102	P002	2
0103	P003	5
0104	P004	1
0105	P005	4

For the given instance, the `Quantity` values need to be updated by decreasing 1 for those entries where the current values are greater than 2. What is the correct SQL Query for updating the current instance?

MODIFY used with ALTER

- a) `MODIFY OrderDetails Quantity=Quantity-1 where Quantity>2;`
- b) `UPDATE OrderDetails set Quantity=Quantity-1 where Quantity>2;`
- c) `UPDATE OrderDetails Quantity=Quantity-1 where Quantity>2;`
- d) `ALTER OrderDetails set Quantity=Quantity-1 where Quantity>2;`

Answer: b)

Explanation: To modify an existing value in a table, the `UPDATE` statement is used in SQL, followed by setting the new values for the columns based on a specified condition. The correct query is:

```
UPDATE Tablename
SET column1 = value1, column2 = value2, column3 = value3,...
WHERE condition;
```

Thus, option (b) is the correct answer.

Question 2

Consider the following two tables representing a book store system:

OrderDetails		
OrderID	BookID	Quantity
0101	B001	5
0102	B003	2
0103	B002	7
0104	B001	1

BookDetails	
BookID	Category
B001	Fiction
B002	Science
B003	Fiction

What is the output of the following SQL query?

```
SELECT Category, SUM(Quantity) FROM OrderDetails, BookDetails WHERE OrderDetails.BookID = BookDetails.BookID GROUP BY Category;
```

- a) Fiction: 8, Science: 7
- b) Fiction: 6, Science: 7
- c) Fiction: 9, Science: 7
- d) Fiction: 8, Science: 6

Answer: a)

Explanation: The query performs an inner join between `OrderDetails` and `BookDetails` on `BookID`. The result is grouped by `Category`, and the `SUM(Quantity)` is calculated for each category. The intermediate table after the join is:

OrderID	BookID	Quantity	Category
0101	B001	5	Fiction
0102	B003	2	Fiction
0103	B002	7	Science
0104	B001	1	Fiction

The total quantities for each category are calculated as: Fiction: $5 + 2 + 1 = 8$; Science: 7. Thus, option (a) is correct.

```
CREATE TABLE Person(Id int NOT NULL, LastName varchar(255) NOT NULL, FirstName varchar(255), Age int, CONSTRAINT PK_Person PRIMARY KEY (ID, LastName));
```

You can insert records as long as the combination of `Id` and `LastName` is unique:

```
sql Copy Edit

INSERT INTO Person VALUES (1, 'Smith', 'John', 30);
INSERT INTO Person VALUES (2, 'Johnson', 'Alice', 25);
INSERT INTO Person VALUES (1, 'Doe', 'Jane', 28); -- Valid since (Id, LastName) is unique
```

Id	LastName	FirstName	Age
1	Smith	John	30
2	Johnson	Alice	25
1	Doe	Jane	28

Question 3

Consider the following instance ProductDetails of an e-commerce application:

ProductDetails			
ProductID	Name	Price	Category
P001	Laptop	50000	Electronics
P002	Mobile	20000	Electronics
P003	Shoes	3000	Footwear
P004	Watch	1500	Accessories
P005	Shirt	800	Clothing

If you update the underlying table, the view will reflect the changes:

Identify the correct statement to create a VIEW on the ProductDetails table to retrieve Name and Price of all products in the Electronics category, and name the view as 'Electronics.Products':

- a) Create Electronics_Products
AS SELECT Name, Price
FROM ProductDetails
WHERE Category = 'Electronics';
- b) Create view Electronics_Products
AS SELECT Name, Price
FROM ProductDetails
WHERE Category = 'Electronics';
- c) Create view Electronics_Products
ON ProductDetails
SELECT Name, Price
WHERE Category = 'Electronics';
- d) Create view Electronics_Products
TO ProductDetails
SELECT Name, Price
WHERE Category = 'Electronics';

Answer: b)

Explanation: The correct syntax for creating a VIEW is:

```
CREATE VIEW viewname  
AS SELECT column1, column2, ...  
FROM tablename  
WHERE condition;
```

So, option (b) is correct.

Question 4

Consider the following table for a ride-sharing application:

DriverDetails	
DriverID	Location
D001	Bangalore
D002	Hyderabad
D003	Chennai
D004	Pune
D005	Bhopal
D006	Delhi

Which of the following options will be present in the output generated by the SQL query:

```
SELECT Location FROM DriverDetails WHERE Location LIKE 'B%' AND Location LIKE '%e' ;
```

- a) Bangalore
- b) Bhopal
- c) Pune
- d) Hyderabad

Answer: a)

Explanation: The operator 'LIKE' uses patterns described by special characters: percent (%) and underscore (_). The % character matches any substring, and the _ character matches any single character.

The clause 'WHERE Location LIKE 'B%' matches locations starting with 'B', such as 'Bangalore' and 'Bhopal'.

The clause "AND Location LIKE '%e' " further filters the result to include locations ending with 'e'. Thus, only 'Bangalore' satisfies both conditions.

Question 5

Consider the two instances:

OrderDetails		
OrderID	ProductID	Quantity
0001	P001	10
0002	P002	5
0003	P003	8
0004	P002	3

ProductDetails	
ProductID	ProductName
P001	Laptop
P002	Mouse
P003	Keyboard

Which of the following relational algebra operations will generate the following output:

OrderID	ProductID	Quantity	ProductName
0001	P001	10	Laptop
0002	P002	5	Mouse
0003	P003	8	Keyboard
0004	P002	3	Mouse

- a) OrderDetails NATURAL JOIN ProductDetails
- b) OrderDetails LEFT OUTER JOIN ProductDetails
- c) OrderDetails RIGHT OUTER JOIN ProductDetails
- d) OrderDetails EQUI JOIN ProductDetails ON OrderDetails.ProductID=ProductDetails.ProductID

Answer: a)

Explanation: A NATURAL JOIN joins two relations based on all the columns with the same name. In this case, the common column is ProductID. The result includes rows where there is a matching ProductID in both relations.

An EQUI JOIN is a join where the join condition contains an equality operator. An EQUI JOIN returns only the rows that have equal values for the specified column(s) and the compared column(s) twice. Hence, option (a) is the correct answer.

Question 6

Which of the following statements is **incorrect**?

- a) The **INSERT** command is used to remove\modify rows in a relation.
- b) The **UPDATE** command is used to modify data (values in rows) of a relation.
- c) The **DELETE** command is used to remove all data from a relation.
- d) The **DELETE** command is used to remove a relation entirely.

Answer: a), d)

Explanation: The **INSERT** command is used to add rows, not modify or remove them. The **DELETE** command removes data (rows) from a relation but does not remove the relation itself. Hence, options (a) and (d) are the answer.

Question 7

Consider the following instance of the DriverDetails(DriverID, Name, Age, City) relation.

DriverDetails			
DriverID	Name	Age	City
D001	John	25	New York
D002	Sarah	32	Los Angeles
D003	Michael	30	Chicago
D004	Alice	28	New York
D005	David	35	Los Angeles
D006	Emma	40	Chicago
D007	Olivia	22	New York
D008	James	28	Los Angeles

What will be the output of the following query?

```
SELECT DriverID, Name, City
FROM DriverDetails d1
WHERE Age > (
    SELECT AVG(Age)
    FROM DriverDetails d2
    WHERE d1.City = d2.City
);
```

a)

DriverID	Name	City
D002	Sarah	Los Angeles
D004	Alice	New York
D005	David	Los Angeles
D006	Emma	Chicago

b)

DriverID	Name	City
D003	Michael	Chicago
D004	Alice	New York
D008	James	Los Angeles

c)

DriverID	Name	City
D001	John	New York
D002	Sarah	Los Angeles
D003	Michael	Chicago
D007	Olivia	New York

d)

DriverID	Name	City
D002	Sarah	Los Angeles
D005	David	Los Angeles
D006	Emma	Chicago

Answer: a)

Explanation: The SQL query selects drivers whose age is greater than the average age of other drivers in the same city. The subquery calculates the average age of drivers in the same city for each driver (using 'd1.City = d2.City'), and the outer query selects the drivers whose age is greater than the calculated average for their city.

Hence, option a) is correct.

Question 8

Consider the following instance of the DriverDetails(DriverID, Name, Age, City) relation.

DriverDetails			
DriverID	Name	Age	City
D001	John	25	New York
D002	Sarah	32	Los Angeles
D003	Michael	30	Chicago
D004	Alice	28	New York
D005	David	35	Los Angeles
D006	Emma	40	Chicago
D007	Olivia	22	New York
D008	James	28	Los Angeles

What will be the output of the following query?

```
SELECT Name, Age, City
FROM DriverDetails
WHERE Age > (
    SELECT AVG(Age)
    FROM DriverDetails e2
    WHERE e2.City = 'Los Angeles');
```

a)

Name	Age	City
Sarah	32	Los Angeles
David	35	Los Angeles

b)

Name	Age	City
John	25	New York
Alice	28	New York
Olivia	22	New York

c)

Name	Age	City
Sarah	32	Los Angeles
David	35	Los Angeles
Emma	40	Chicago

d)

Name	Age	City
Sarah	32	Los Angeles
David	35	Los Angeles
Michael	30	Chicago
Alice	28	New York

Answer: c)

Explanation: The SQL query selects drivers whose age is greater than the average age of drivers in the 'Los Angeles' city. The subquery calculates the average age of drivers in 'Los Angeles', which is $(32+35+28)/4 = 31.67$. The outer query selects drivers whose age is greater than 31.67.

Therefore, the correct answer is option c).

Question 9

Consider the following instance `DriverDetails` of a transportation application:

DriverDetails			
DriverID	Name	Age	City
D001	John	25	New York
D002	Sarah	32	Los Angeles
D003	Michael	30	Chicago
D004	Alice	28	New York
D005	David	35	Los Angeles
D006	Emma	40	Chicago
D007	Olivia	22	New York
D008	James	28	Los Angeles

What will be the output of the following query?

```
SELECT * FROM DriverDetails
WHERE City IN ('New York', 'Chicago');
```

a)

DriverID	Name	Age	City
D001	John	25	New York
D003	Michael	30	Chicago
D004	Alice	28	New York
D006	Emma	40	Chicago
D007	Olivia	22	New York

b)

DriverID	Name	Age	City
D002	Sarah	32	Los Angeles
D005	David	35	Los Angeles

c)

DriverID	Name	Age	City
D001	John	25	New York
D002	Sarah	32	Los Angeles
D003	Michael	30	Chicago

d)

DriverID	Name	Age	City
D005	David	35	Los Angeles
D008	James	28	Los Angeles

Answer: a)

Explanation: The query uses the `IN` operator to select drivers whose `City` is either 'New York' or 'Chicago'. The rows that match these conditions are shown in the output table. Hence, option a) is correct.

Question 10

Consider the following instance `DriverDetails` of a transportation application:

DriverDetails			
DriverID	Name	Age	City
D001	John	25	New York
D002	Sarah	32	Los Angeles
D003	Michael	30	Chicago
D004	Alice	28	New York
D005	David	35	Los Angeles
D006	Emma	40	Chicago
D007	Olivia	22	New York
D008	James	28	Los Angeles

Identify the correct statement to find the `DriverID`, `Name`, and `City` of `DriverDetails` table whose `Age` is between 25 and 35.

- a) `SELECT DriverID, Name, City`
`FROM DriverDetails`
`WHERE Age AS (25, 35);`
- b) `SELECT DriverID, Name, City`
`FROM DriverDetails`
`WHERE Age IN (25, 35);`
- c) `SELECT DriverID, Name, City`
`FROM DriverDetails`
`WHERE Age BETWEEN 25 AND 35;`
- d) `SELECT DriverID, Name, City`
`FROM DriverDetails`
`WHERE Age BETWEEN (25, 35);`

Answer: c)

Explanation: The `BETWEEN` operator selects values within a given range. The values can be numbers, text, or dates, and both the start and end values are included.

The correct syntax for using the `BETWEEN` operator is:

```
SELECT columnname(s)
FROM tablename
WHERE columnname BETWEEN value1 AND value2;
```

The `IN` operator allows specifying multiple values in a `WHERE` clause. It is a shorthand for multiple `OR` conditions, but it is not used for ranges.

Hence, option c) is correct.



Course Name: Database Management System

Assignment 3 - Week 3 (Jan 2025)

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10

Total mark: 10 X 2 = 20

Question 1

An organization, collecting car renters' information, considers the following relations:

Renter(Renter_ID, Contact)

Renting(Renter_ID, Car_Number)

Car(Car_Number, Model).

What will the following relational algebra expression return to the organization?

$$\Pi_{\text{Contact}}(\text{Renter}) - \Pi_{\text{Contact}}(\text{Renter} \bowtie \text{Renting})$$

- a) The Contacts of those Renters who are Renting at most one Car at that instance.
- b) The Contacts of those Renters who are Renting every Car at that instance.
- c) The Contacts of those Renters who are Renting at least one Car at that instance.
- d) The Contacts of those Renters who are not Renting any Car at that instance.

Answer: d)

Explanation: $\Pi_{\text{Contact}}(\text{Renter})$ returns all the Contacts from the enlisted Renters.

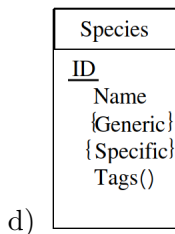
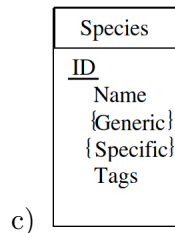
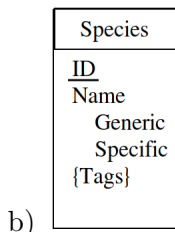
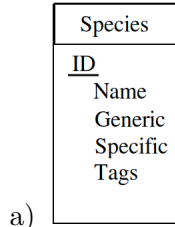
$\Pi_{\text{Contact}}(\text{Renter} \bowtie \text{Renting})$ returns the Contacts of those Renters whose IDs have matched with some ID in Renting i.e. the Contacts of those Renters who are Renting some Car at that instance.

Therefore, the (Contacts of all Renters – the Contacts of those Renters renting some Car) gives the Contacts of those Renters who are **not** Renting any Car.

Hence, option d) is correct.

Question 2

An institute, maintaining a biological information system, designs an entity **Species** having the attributes **ID**, **Name**, and **Tags**. **ID** is unique and cannot be NULL. The **Name** is composed of **Generic** and **Specific** names. **Tags** can have multiple values. Which of the following ER notation represents the entity **Species** correctly?



Answer: b)

Explanation: From the given problem, the **ID** should be the key attribute, denoted by underline, **Name** should be a composite attribute where the sub-attributes **Generic**, **Specific** should be denoted with an indentation and **Tags** is a multivalued attribute enclosed by **{}**. Hence, option b) is correct.

Question 3

Consider 2 instances:

Chapter		
PageNo	Topic	Length
1	Introduction	5
2	SQL	6
3	RA	7
4	FD	8
3	ERD	2

Paragraph	
PageNo	Topic
2	SQL
3	RA

Which of the following operations return the 2 tuples that are present in Paragraph?

- a) $(\text{Chapter}) - (\Pi_{\text{PageNo}, \text{Topic}}(\text{Chapter}) - \text{Paragraph})$
- b) $(\text{Chapter} \times \text{Paragraph}) \cap (\text{Paragraph} \times \text{Chapter})$
- c) $(\text{Paragraph}) \cap (\Pi_{\text{PageNo}, \text{Topic}}(\text{Chapter}) \cup \text{Paragraph})$
- d) $\Pi_{\text{PageNo}, \text{Topic}}((\Pi_{\text{PageNo}, \text{Topic}}(\text{Chapter}) - \text{Paragraph}) \times \text{Paragraph})$

Answer: c)

Explanation: Option (a) results in invalid – operation as the “arity” of Chapter and $(\Pi_{\text{PageNo}, \text{Topic}}(\text{Chapter}) - \text{Paragraph})$ are not same.

Option (b) is incorrect as the \cap operation has relations with incompatible domains of attributes on L.H.S and R.H.S.

Option (c) is correct. As $\text{Paragraph} \subset \text{Chapter}$, $\Pi_{\text{PageNo}, \text{Topic}}(\text{Chapter}) \cup \text{Paragraph}$ results in all the values of PageNo, Topic in Chapter and the \cap selects the ones in Paragraph only.

Option (d) is incorrect as the outermost projection operation results in ambiguity.

Question 4

Consider the relations Temp_India(Place, Celsius) and Temp_France(Place, Celsius) and the following instances:

Temp_India		Temp_France	
Place	Celsius	Place	Celsius
Ahmedabad	29	Nice	20
Jaipur	22	Paris	18
Roorkee	22	Dijon	11
Kolkata	21		

$X = \text{Temp_India} - ((\text{Temp_India} \cup \text{Temp_France}) - (\text{Temp_India} - \text{Temp_France}))$

Which of the following is X?

- a) Temp_India
- b) Temp_France
- c) $\text{Temp_India} \cup \text{Temp_France}$
- d) $\text{Temp_India} \cap \text{Temp_France}$

Answer: a)

Explanation: $((\text{Temp_India} \cup \text{Temp_France}))$ produces all tuples from both relations.

$(\text{Temp_India} - \text{Temp_France}) = \text{Temp_India}$.

Thus, $((\text{Temp_India} \cup \text{Temp_France}) - (\text{Temp_India} - \text{Temp_France})) = \text{Temp_France}$

and $\text{Temp_India} - ((\text{Temp_India} \cup \text{Temp_France}) - (\text{Temp_India} - \text{Temp_France})) = \text{Temp_India}$.

Hence, option (a) is correct.

Question 5

Consider the relational schema `Photography(Photographer, Camera, Theme, Location)`. Choose the correct Tuple Relational Calculus that represents the following statement
“Display all photographers who clicked photographs with a DSLR camera in Darjeeling.”

- a) $\{p, t \mid p, t \in (t[\text{Photographer}] = p[\text{Photographer}] \wedge p[\text{Camera}] = \text{'DSLR'} \wedge p[\text{Location}] = \text{'Darjeeling'})\}$
- b) $\{t \mid p < p[\text{Camera}] = \text{'DSLR'} \wedge p[\text{Location}] = \text{'Darjeeling'}\}$
- c) $\{p, t \mid \exists p \in \text{Photography } (p[\text{Camera}] = \text{'DSLR'} \wedge p[\text{Location}] = \text{'Darjeeling'})\}$
- d) $\{t \mid \exists p \in \text{Photography } (t[\text{Photographer}] = p[\text{Photographer}] \wedge p[\text{Camera}] = \text{'DSLR'} \wedge p[\text{Location}] = \text{'Darjeeling'})\}$

Answer: d)

Explanation: The tuple to be selected is represented by ‘*t*’ and the selection conditions are written with \wedge as per the given question. According to the projection and selection operations shown in lecture slides 12.26 - 12.28, option (d) shows the correct syntax and semantics of the Tuple Relational Calculus.

Question 6

Consider the following instance of the relation $\text{CakeShop}(\text{Owner}, \text{Name})$:

CakeShop	
Owner	Name
Peter	BakeCake
Viola	CakeStop
John	Cherry
Peter	SweetTooth
Peter	FudgePalace
Viola	Desserted

Which of the following relational algebra expression does generate names of those **Owners** who owns two or more **CakeShops**?

- a) $\Pi_{C1.Owner}(\sigma_{C1.Owner \neq C2.Owner \wedge C1.Name \neq C2.Name}(\rho_{C1}(\text{CakeShop}) \times \rho_{C2}(\text{CakeShop})))$
- b) $\Pi_{C1.Owner}(\sigma_{C1.Owner = C2.Owner \wedge C1.Name \neq C2.Name}(\rho_{C1}(\text{CakeShop}) \times \rho_{C2}(\text{CakeShop})))$
- c) $\Pi_{C1.Owner}(\sigma_{C1.Owner = C2.Owner \wedge C1.Name = C2.Name}(\rho_{C1}(\text{CakeShop}) \times \rho_{C2}(\text{CakeShop})))$
- d) $\Pi_{C1.Owner}(\sigma_{C1.Owner \neq C2.Owner \wedge C1.Name = C2.Name}(\rho_{C1}(\text{CakeShop}) \times \rho_{C2}(\text{CakeShop})))$

Answer: b)

Explanation: The Cartesian product between the renamed tables will generate 36 tuples and the condition $\sigma_{C1.Owner = C2.Owner \wedge C1.Name \neq C2.Name}$ selects the following tuples out of the 36 tuples:

C1.Owner	C1.Name	C2.Owner	C2.Name
Peter	BakeCake	Peter	SweetTooth
Peter	BakeCake	Peter	FudgePalace
Viola	CakeStop	Viola	Desserted
Peter	SweetTooth	Peter	BakeCake
Peter	SweetTooth	Peter	FudgePalace
Peter	FudgePalace	Peter	BakeCake
Peter	FudgePalace	Peter	SweetTooth
Viola	Desserted	Viola	CakeStop

The final projection of $C1.Owner$ will only produce $\{\text{Peter}, \text{Viola}\}$.

Hence, option (b) is correct.

Question 7

Consider the relational schema `FormulaBook(Page, Sl, Formula)`.

An embedded SQL query is written with C as host language to derive some `Formula` entered in the `FormulaBook` according to the following conditions

```
EXEC SQL
DECLARE c CURSOR FOR
SELECT Formula
FROM FormulaBook
WHERE Page> :PageGiven AND Sl=:SLGiven
END EXEC
```

Select the correct option for declaring the variables `PageGiven` and `SLGiven`.

- a) EXEC-SQL BEGIN
 int :PageGiven ;
 int :SLGiven ;
EXEC-SQL END;
- b) EXEC-SQL BEGIN DECLARE SECTION
 int PageGiven ;
 int SLGiven ;
EXEC-SQL END DECLARE SECTION;
- c) EXEC-SQL BEGIN DECLARE SECTION
 int :PageGiven ;
 int :SLGiven ;
EXEC-SQL END DECLARE SECTION;
- d) EXEC-SQL BEGIN
 int PageGiven ;
 int SLGiven ;
EXEC-SQL END;

Answer: b)

Explanation: The declaration of the variables are enclosed within

```
EXEC-SQL BEGIN DECLARE SECTION
EXEC-SQL END DECLARE SECTION;
```

The variable declaration follow the host language syntax. Hence, option (b) is correct.

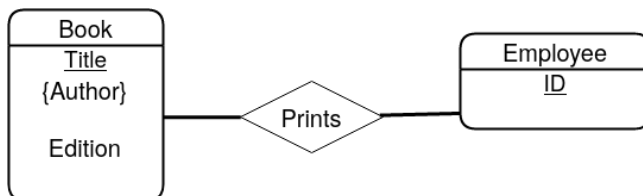
Question 8

In a printing company, a **Book** is printed by multiple **Employees** and an **Employee** prints multiple **Books**. The **Books** are identified by their unique **Title**. A **Book** also has an **Edition** and multiple values of the attribute **Author**. An **Employee** has a unique **ID**. Assuming there is a **Prints** relation between **Book** and **Employee**, what will be the correct schema for **Prints** and **Book**?

- a) **Prints**(Title, ID)
 Book(Title, Edition, Author)
- b) **Prints**(ID)
 Book(Title Edition)
 Book.author(Title, Author)
- c) **Prints**(Title, ID)
 Book(Title, Edition)
 Book.author(Title, Author)
- d) **Prints**(Title, ID, Author)
 Book(Title, Edition)

Answer: c)

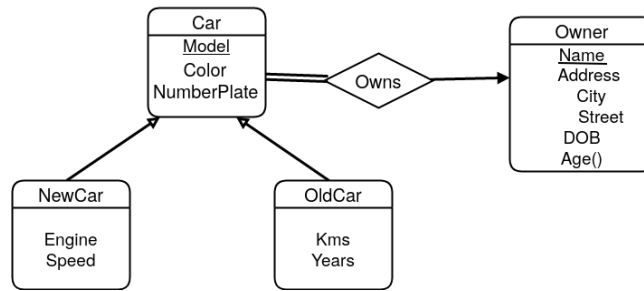
Explanation: The following represents the ER diagram of the given case.



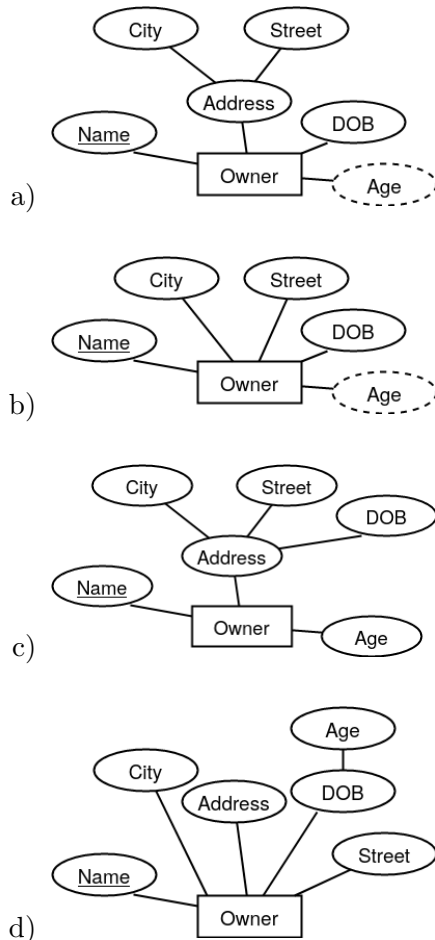
In many-to-many relation, the relation **Prints** should have the primary keys of the participating entities and since **Author** is multivalued, a separate schema has to be created named **Book_author**. Hence, option (c) is correct.

Question 9

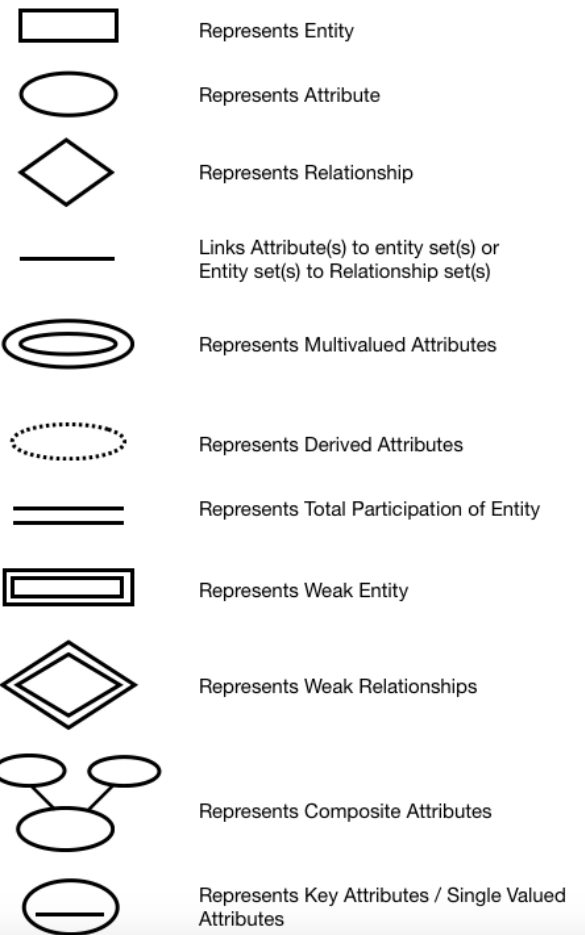
Consider the Entity Relationship Diagram:



Which of the following is an alternate ER notation for **Owner**?



Age is derived

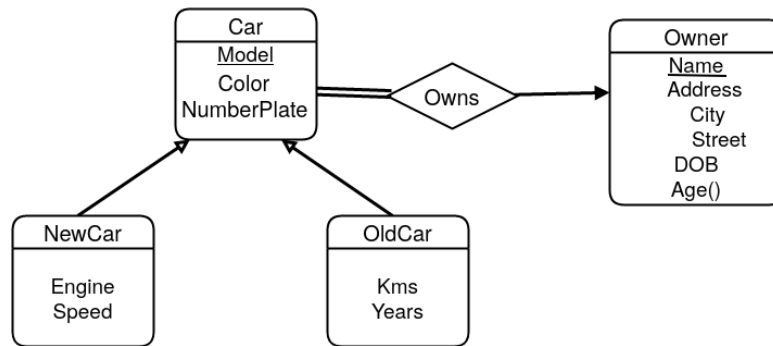


Answer: a)

Explanation: As per the alternate ER notation shown in lecture slide 15.27.

Question 10

Consider the Entity Relationship Diagram:



Which of the following is/are not true?

- a) NewCar and OldCar are Overlapping specializations of Car.
- b) Participation of Car in Owns is total.
- c) Age is a composite attribute of Owner.
- d) DOB is a component attribute of Address.

Answer: c), d)

Explanation: Option (a) and (b) are true as per the ER Diagram notations. Age is a derived attribute of Owner. DOB is not a component attribute of Address. Hence, options (c) and (d) are not true.



Course Name: Database Management System

Assignment 4 - Week 4 (Jan 2025)

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10

Total mark: $10 \times 2 = 20$

Question 1

Consider the relation MchineParts(PartNo, MachineNo, Color, Shape, Position) with the following Functional Dependencies:

FD1: $\text{PartNo} \rightarrow \{\text{MachineNo}, \text{Color}\}$

FD2: $\text{MachineNo} \rightarrow \text{Shape}$

FD3: $\{\text{Color}, \text{Shape}\} \rightarrow \text{Position}$

FD4: $\text{Position} \rightarrow \text{PartNo}$

Which of the following is not a candidate key?

- a) {Color, Shape}
- b) {MachineNo, Shape}
- c) {MachineNo, Color}
- d) Position

Answer: b)

Explanation: The closure $(\{\text{MachineNo}, \text{Shape}\})^+ = \{\text{MachineNo}, \text{Shape}\}$. The closure of all other options produce the attributes {PartNo, MachineNo, Color, Shape, Position}. Hence, option (b) is correct.

Question 2

Consider the following relational schema: `Airway(Flno, Flname, Source, Destination)`.
The following functional dependencies hold:

FD1: `Flname` \rightarrow `Source`

FD2: `Source` \rightarrow `Flno`

FD3: `Source` \rightarrow `Destination`

What is the highest normal form of `Airway`?

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

Answer: b)

Explanation: `Flname` is the key. The relation is in 1NF.

Since there is no partial dependency, it is in 2NF.

But, FD1, FD2 and FD1, FD3 show transitive dependency. So it is not in 3NF.

Hence, option b) is correct.

Question 3

Consider the relational schema $\text{Investor}(\text{InvestmentNo}, \text{Amount}, \text{Tenure})$. Two organisations identify 2 different sets of Functional Dependencies for Investor :

$\text{OrganisationA} = \{$

FD1: $\text{InvestmentNo} \rightarrow \{\text{Amount}, \text{Tenure}\}$

FD2: $\text{Amount} \rightarrow \text{InvestmentNo}$

FD3: $\text{Tenure} \rightarrow \text{InvestmentNo}$

$\}$

$\text{OrganisationB} = \{$

FD1: $\text{InvestmentNo} \rightarrow \text{Amount}$

FD2: $\text{Amount} \rightarrow \text{Tenure}$

FD3: $\text{Tenure} \rightarrow \text{InvestmentNo}$

$\}$

Which of the following statements is(are) true?

- a) Only OrganisationA covers OrganisationB .
- b) Only OrganisationB covers OrganisationA .
- c) OrganisationA and OrganisationB are equivalent.
- d) OrganisationA and OrganisationB cannot be compared.

Answer: c)

Explanation:

Step 1:

The closures of attribute sets to the L.H.S of FD1, FD2, FD3 under OrganisationA are estimated using the FDSs in OrganisationA and OrganisationB .

$\text{OrganisationA: } (\text{InvestmentNo})^+ = \{\text{InvestmentNo}, \text{Amount}, \text{Tenure}\}$

$\text{OrganisationB: } (\text{InvestmentNo})^+ = \{\text{InvestmentNo}, \text{Amount}, \text{Tenure}\}$

$\text{OrganisationA: } (\text{Amount})^+ = \{\text{InvestmentNo}, \text{Amount}, \text{Tenure}\}$

$\text{OrganisationB: } (\text{Amount})^+ = \{\text{InvestmentNo}, \text{Amount}, \text{Tenure}\}$

$\text{OrganisationA: } (\text{Tenure})^+ = \{\text{InvestmentNo}, \text{Amount}, \text{Tenure}\}$

$\text{OrganisationB: } (\text{Tenure})^+ = \{\text{InvestmentNo}, \text{Amount}, \text{Tenure}\}$

Hence, OrganisationA covers OrganisationB .

Step 2:

The closures of attribute sets to the L.H.S of FD1, FD2, FD3 under OrganisationB are estimated using the FDSs in OrganisationA and OrganisationB . Since there is no additional attribute set to the L.H.S of any FD in OrganisationB , from step 1 it can be concluded that OrganisationB covers OrganisationA .

Step 3:

Since, OrganisationA covers OrganisationB and OrganisationB covers OrganisationA , OrganisationA and OrganisationB are equivalent. Hence, option (c) is correct.

Question 4

Consider the following relation Work:

Work		
Incharge	Department	Experience
S. Roy	Production	5
A. Bera	Sales	2
S. Roy	HR	5
S. Rai	Development	3
B. Mallik	Testing	4
S. Sinha	Testing	4

Which of the following functional dependencies can hold on Work?

- a) $\text{Department} \rightarrow \text{Experience}$
- b) $\text{Incharge} \rightarrow \text{Experience}$
- c) $\text{Experience} \rightarrow \text{Incharge}$
- d) $\text{Incharge} \rightarrow \text{Department}$

Answer: a), b)

Explanation: For options a), b), the attributes on the L.H.S can uniquely identify the attributes on the R.H.S.

Question 5

In a relation `BloodDonation`, the attributes are: `DonorID`, `DonorName`, `DonorAddress`, and `DonorBloodGroup`. `DonorID` can uniquely identify `DonorName`, `DonorAddress`. `DonorName` can uniquely identify `DonorBloodGroup`. If the first donor, having `DonorID = A1`, `DonorName = Aman`, `DonorAddress = Kolkata` and `DonorBloodGroup = A+` was registered and recorded under the schema, which of the following records will be invalid as another entry in `BloodDonation`?

- a) Donor having `DonorID=A2`, `DonorName=Aman`, `DonorAddress=Mumbai`, `DonorBloodGroup=B+`.
- b) Donor having `DonorID=A2`, `DonorName=Aman`, `DonorAddress=Kolkata`, `DonorBloodGroup=A+`.
- c) Donor having `DonorID=A2`, `DonorName=Nikhil`, `DonorAddress=Mumbai`, `DonorBloodGroup=B+`.
- d) Donor having `DonorID=A2`, `DonorName=Nikhil`, `DonorAddress=Mumbai`, `DonorBloodGroup=A+`.

Answer: a)

Explanation: If the other donor has `DonorID = A2`, `DonorName = Aman`, `DonorAddress = Mumbai` and `DonorBloodGroup= B+`, it violates the functional dependency `DonorName → DonorBloodGroup`. This is because, `Aman` will be associated with both `A+` and `B+`. Hence, the record in option (a) cannot be the other entry in the instance. Hence, option (a) is correct.

Question 6

A software company maintains a relation **OnlineRace** with attributes like **GameID**, **Player**, **CarType**, **CarColor**, **RaceTime**, **Laps**. **GameID** can uniquely identify all other attributes. **Player** can uniquely identify **CarType** and **Laps** is identified by **RaceTime**. What is the highest normal form of **OnlineRace**?

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

Answer: b)

Explanation: The relation **OnlineRace** has the following functional dependencies:

FD1: $\text{GameID} \rightarrow \{\text{Player}, \text{CarType}, \text{CarColor}, \text{RaceTime}, \text{Laps}\}$

FD2: $\text{Player} \rightarrow \text{CarType}$

FD3: $\text{RaceTime} \rightarrow \text{Laps}$

GameID is the key and there are no partial dependencies. However, FD2 and FD3 show transitive dependencies as they are dependencies between non-key attributes.

Hence, the highest Normal Form is 2NF.

Question 7

Consider the relation `PollutionControl(FactoryID, EmissionLevel, EmissionEffect, Locality)` with the following Functional Dependencies:

FD1: $\{FactoryID, EmissionLevel\} \rightarrow \{EmissionEffect, Locality\}$

FD2: $\{EmissionLevel, EmissionEffect\} \rightarrow Locality$

`PollutionControl` is decomposed into the following:

`PollutionControl1(FactoryID, EmissionLevel, EmissionEffect)` and

`PollutionControl2(FactoryID, EmissionLevel, Locality)`

Which of the following statements is(are) true?

- a) The decomposition is lossless but not dependency preserving.
- b) The decomposition is dependency preserving but not lossless.
- c) The decomposition is both lossless and dependency preserving.
- d) The decomposition is neither lossless nor dependency preserving.

Answer: a)

Explanation: The decomposition is lossless as

(1) $PollutionControl1 \cap PollutionControl2 = \{FactoryID, EmissionLevel\}$

(2) $PollutionControl1 \cup PollutionControl2 = \{FactoryID, EmissionLevel, EmissionEffect, Locality\}$

(3) $\{FactoryID, EmissionLevel\}$ is a primary key. Hence $\{FactoryID, EmissionLevel\} \rightarrow PollutionControl1, PollutionControl2$.

The decomposition is not dependency preserving as

(1) Non-trivial dependencies derived from original FDs for `PollutionControl1` are $Set1 = \{ \{FactoryID, EmissionLevel\} \rightarrow EmissionEffect \}$

(2) Non-trivial dependencies derived from original FDs for `PollutionControl2` are $Set2 = \{ \{FactoryID, EmissionLevel\} \rightarrow Locality \}$

(3) FDs derivable from $Set1 \cup Set2 = \{ \{FactoryID, EmissionLevel\} \rightarrow \{EmissionEffect, Locality\} \}$ which is FD1. However, FD2 cannot be derived. Hence option (a) is correct.

Question 8

Consider the relation $\text{Tea}(\text{Mfd}, \text{Origin}, \text{Type}, \text{Flavor}, \text{Rating})$ with the following Functional Dependencies:

FD1: $\text{Mfd} \rightarrow \{\text{Origin}, \text{Type}\}$

FD2: $\text{Origin} \rightarrow \text{Flavor}$

FD3: $\{\text{Type}, \text{Flavor}\} \rightarrow \text{Rating}$

FD4: $\text{Rating} \rightarrow \text{Mfd}$

What is the maximum number of candidate keys for the relation Tea ?

a) 2

b) 3

c) 4

d) 5

Answer: c)

Explanation: By estimating the closure of all combination of the attributes, it can be observed that the closure of the following attributes produce all other attributes:

$\text{Mfd}, \text{Rating}, \{\text{Origin}, \text{Type}\}, \{\text{Type}, \text{Flavor}\}.$

Hence, these 4 are the candidate keys.

Question 9

Consider the relation TimeCounter(LogID, Days, Hours, Mins, Secs) with the following functional dependencies:

FD1: $\text{LogID} \rightarrow \{\text{Days}, \text{Hours}\}$

FD2: $\text{Days} \rightarrow \text{Mins}$

FD3: $\{\text{Hours}, \text{Mins}\} \rightarrow \text{Secs}$

FD4: $\text{Secs} \rightarrow \text{LogID}$

Which of the following functional dependencies violate the BCNF form?

a) FD1

b) FD2

c) FD3

d) FD4

Answer: b)

Explanation: From the given functional dependencies, the candidate keys can be identified as LogID, Sec, {Days, Hours} and {Hours, Mins}. FD2 is neither trivial nor the L.H.S of this dependency form a superkey of the given relation and hence, violates BCNF form. Hence, options (b) is correct.

Question 10

Consider the relation `Carnival(CID, Location, Span, RideCount, ShowCount, StartTime, EndTime, Crowd)` with the following functional dependencies:

FD1: $CID \rightarrow \{Location, Span, RideCount\}$

FD2: $\{CID, RideCount\} \rightarrow ShowCount$

FD3: $\{ShowCount, StartTime, EndTime\} \rightarrow Crowd$

FD4: $StartTime \rightarrow \{EndTime, Crowd\}$

Which of these functional dependencies can be eliminated without changing the primary key of the relation?

a) FD1

b) FD2

c) FD3

d) FD4

Answer: c)

Explanation: The given functional dependencies imply that only attributes `CID` and `StartTime` cannot be reached from any other attributes. Hence, the key for this relation can be evaluated to be $\{CID, StartTime\}$. If we remove any other FD except FD3, the key of the relation will change. Hence, option (c) is correct.



NPTEL Online Certification Courses

Indian Institute of Technology Kharagpur
Jan 2025



Course Name: Database Management System

Assignment 5 - Week 5 (Jan 2025)

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10

Total mark: $10 \times 2 = 20$

Question 1

Which of the followings can be the solution(s) of Password Leakage in Database Servers?

- a) Storing encrypted passwords in database and in scripts.
- b) Reusing passwords across different sites.
- c) Single factor authentication using passwords.
- d) Two-factor authentication (password plus one-time password sent by SMS).

Answer: a), d)

Explanation: Refer to lecture slide 22.

Question 2

Which of the following statement(s) is/are incorrect?

- a) Javascript is a Client-Side Scripting language.
- b) World Wide Web is a language used for designing web pages.
- c) Hyper Text Transfer Protocol is a connection-oriented protocol.
- d) J2EE platform supports parallel processing.

Answer: b), c)

Explanation: World Wide Web is distributed information system based on hypertext. Hyper Text Transfer Protocol is connectionless. So, these two options are wrong.

Refer to Module 21.

So, options b) and c) are the answer.

Question 3

Consider the following table DEFAULTER (ID, NAME, AMOUNT). If we want to create an index on DID column, which type of indexing will be preferred?

DEFAULTER		
DID	NAME	AMOUNT
415	AISHA	20000
124	JOHN	5000
347	EDWARD	100
145	RAY	200
209	AKASH	900

- a) Sparse index
- b) Secondary index
- c) Dense index
- d) Clustering index

Answer: c)

Explanation: When the file is not sorted on the indexed field or when the index file is small, compared to the size of the memory, it is preferable to use dense indexing. The above table has only 5 records and **DID** field is not in sorted order.

Hence, option c) is correct.

Question 4

A magnetic disk has 16 platter surfaces, 256 tracks per surface, 512 sectors per track, and 1024 bytes per sector. How many number of bits will be required to address a sector?

- a) 17
- b) 21
- c) 27
- d) 31

Answer: b)

Explanation: Number of platter surfaces = 16

Number of tracks per surface = 256

Number of sectors per track = 512

Therefore, total number of sectors

= Total number of platter surfaces x Number of tracks per surface x Number of sectors per track

= $16 \times 256 \times 512$ sectors

= 2^{4+8+9} sectors

= 2^{21}

So, Number of required bits to address the sector = 21 bits

Hence, option b) is correct.

Question 5

Availability of Redundant Arrays of Independent Disks system is 97%. Mean Time Between Failure (MTBF) is 36 days. What is the approximate Mean Time To Repair (MTTR) of the system?

- a) 10.80 Hours
- b) 26.72 Hours
- c) 34.92 Hours
- d) 88.36 Hours

Answer: b)

Explanation: Mean time between failures is the average time between failures.

Mean time between failures (MTBF) = total available time / number of failures.

Mean time to repair is the average time taken to repair the system.

Mean time to repair(MTTR)= total unavailable time /number of failures

Availability = Total available time /(total available time +total unavailable time)

$$\frac{MTBF}{(MTBF+MTTR)} = 97\%$$

$$\frac{36*24*100}{(36*24+MTTR)} = 97$$

$$36 * 24 + MTTR = \frac{36*24*100}{97}$$

$$MTTR = \frac{(36*24*100)-(36*24*97)}{97} \text{ hours}$$

$$MTTR = 26.72 \text{ hours}$$

Hence, option b) is correct.

Question 6

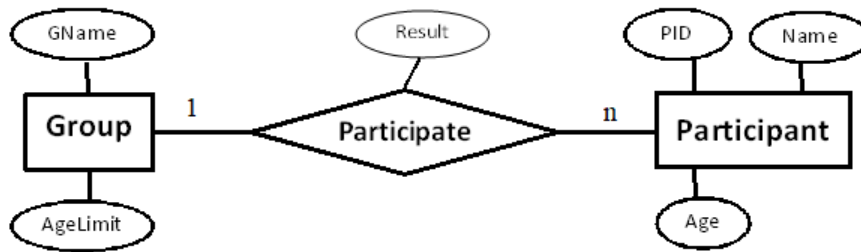
In a Coding Competition, participants individually can enroll their names for the competition. There are many groups in the competition depending on the age of the participants. A participant can participate in only one group and a group can consist of multiple participants. Each group has a unique name and a participant also has a unique id. The result of each group will be maintained separately. Which of the following statement(s) are correct?

- a) Entity Group will not have any primary key.
- b) Participate will be a one-to-many relationship between Group and Participants.
- c) Participate will be a many-to-many relationship between Group and Participant.
- d) Gname can be the foreign key of Participate relation between Group and Participant.

Answer: b), d)

Explanation: If we draw the ER diagram of it, we can see there will be a one-to-many relationship between Group and Participant.

Gname is the primary key of the entity Group table. Hence, can be the foreign key of Participate relation between Group and Participant.



So, options b) and d) are the answer.

Question 7

A flash storage system uses the size of a page 4 KB and 32-bit page-address for its operation. Moreover, the flash translation table is stored as an array. If the size of the flash memory is 64 GB, what will be the size of the **flash translation table**?

- a) 64 MB
- b) 32 MB
- c) 24 MB
- d) 16 MB

Answer: a)

Explanation: Size of the flash memory = 64 GB = 64×2^{30} bytes

Size of a page = 4 KB = 4×2^{10} bytes

Number of pages = $\frac{64 \times 2^{30}}{4 \times 2^{10}} = 16 \times 2^{20}$

Size of page address = 32 bits = 4 bytes

Therefore, the size of the flash translation table = $16 \times 2^{20} \times 4$ bytes = 64 MB

Hence, option a) is correct.

Question 8

Consider the following string of reference:

12, 25, 31, 42, 12, 25, 50, 12, 25, 31, 42, 50

Find the number of replacements (where an existing value is replaced by a new value because the buffer is full) incurred using the least recently used (LRU) buffer replacement algorithm with 3 empty buffer frames.

- a) 12
- b) 10
- c) 8
- d) 7

Answer: d)

Explanation:

Buffer			Comments
12			12 is added to empty buffer
12	25		25 is added to empty buffer
12	25	31	31 is added to empty buffer
42	25	31	12 is least recently used. Hence, replaced by 42
42	12	31	25 is least recently used. Hence, replaced by 12
42	12	25	31 is least recently used. Hence, replaced by 25
50	12	25	42 is least recently used. Hence, replaced by 50
50	12	25	No replacement, 12 present
50	12	25	No replacement, 25 present
31	12	25	50 is replaced, as 12, 25 are recently used
31	42	25	12 is replaced, as it is least recently used element
31	42	50	25 is replaced

Hence, altogether 7 replacements have been done. So, option d) is the answer.

Question 9

Consider a file organization, where the size of one record is 48 bytes, the size of the disk block pointer is 10 bytes, and the size of one disk block is 512 bytes. If organization of the file is sequential and unspanned (one block can store only whole records), maximum how many records can be stored in one block?

- a) 12
- b) 11
- c) 10
- d) 9

Answer: c)

Explanation: The file organization is unspanned. So, one block can store N numbers of whole records and one **block pointer**.

One record size(V) = 48 bytes

Disk block size = 512 bytes

Disk block pointer (p) = 10 bytes

Actual Storage size of one block = $(512-10) = 502$ bytes

$N = 502/48 = 10$

So, maximum 10 records can be stored in a block.

Question 10

Identify the correct statement(s) about the following RAID levels?

- a) RAID 0 provides byte-level striping with mirroring.
- b) RAID 1 provides disk mirroring without striping.
- c) RAID 3 provides byte-level striping with dedicated parity checking.
- d) RAID 5 provides Byte-level striping without parity bits.

Answer: b), c)

Explanation: RAID 0 provides block-level striping without parity or mirroring and RAID 5 provides block-level striping with distributed parity. Therefore, these two statements are given in options a) and d) are incorrect.

Hence, options b) and c) are the answers.

0 = data stripping is done (data is divided)

1 = mirroring is used (copy of one data in disk1 and disk 2)

3 = data is stripped with parity bits

5 = data is stripped & parity bits are stored in different disks

6 = data is stripped & 2 parity bits are stored in different disks



NPTEL Online Certification Courses

Indian Institute of Technology Kharagpur
Jan 2025



Course Name: Database Management System

Assignment 6 - Week 6 (Jan 2025)

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10

Total mark: $10 \times 2 = 10$

Question 1

Which of the statement(s) is (are) correct?

- a) B-tree indexing requires less number of nodes than a corresponding B^+ tree.
- b) Non-leaf nodes of B-tree do not contain record pointers.
- c) Size of non-leaf nodes of B^+ tree is larger than the leaf nodes.
- d) Sequential traversing is also possible in B^+ tree.

Answer: a), d)

Explanation: Key value of non-leaf nodes of B-tree does not appear in the leaf node. So they contain record pointers also.

Size of non-leaf nodes of B^+ tree is smaller than leaf nodes as they do not contain any record pointers. So, these two statements are wrong.

Hence, options a) and d) are the answer.

Question 2

Consider the following two table. Identify the correct worst case time Complexity for the given operation and data structure.

Data Structure and operation	Worst case time complexity
1. Delete from a 2-3-4 Trees	a. $O(1)$
2. Search in an ordered array	b. $O(n)$
3. Search in an Unordered List	c. $O(\log n)$
4. Finding minimum in a min-heap	d. $O(n \log n)$

a) 1-c, 2-b, 3-b, 4-a

b) 1-c, 2-c, 3-b, 4-a

c) 1-c, 2-a, 3-b, 4-b

d) 1-a, 2-b, 3-d, 4-a

Answer: b)

Explanation: The worst-case search time complexity for the operations on data structures are:

Data Structure and operation	Worst case time complexity
1. Delete from a 2-3-4 Trees	c. $O(\log n)$
2. Search in an ordered array	c. $O(\log n)$
3. Search in an unordered list	b. $O(n)$
4. Finding minimum in a min-heap	a. $O(1)$

Hence, option b) is correct.

Question 3

Consider a hash table of size $m = 1000$ and the hash function $h(k) = \text{Round}(k \div 7) \% 1000$. Compute the location where the **key k = 97531** will be stored.

- a) 97
- b) 139
- c) 531
- d) 933

Answer: d)

Explanation: the hash function is $h(k) = \text{Round}(k \div 7) \% 1000$

key k = 97531

So, Location = $\text{Round}(97531 \div 7) \% 1000$

= $\text{Round}(13933) \% 1000$

= $13933 \% 1000$

= 933

Hence, option d) is correct.

Question 4

How many maximum numbers of keys, a B-tree of order 6 can store? The height of the tree is 4 and the root node is assumed to be in height zero.

- a) 9330
- b) 7775
- c) 1554
- d) 1295

Answer: b)

Explanation: A B-tree of order m and height h will have the maximum number of keys when all nodes are completely filled.

A height-0, B-tree of order 6 can store 5 keys.

A height-1, B-tree of order 6 can have 6 children. So, can have $6 \times 5 + 5$ keys.

Similarly, maximum number of keys in a B-tree of height $h = (m^{h+1} - 1) / (m - 1)$ keys

Here, $m=6$ and $h=4$. Hence, maximum keys = $(6^{4+1} - 1) / (6 - 1)$

So, $6^5 - 1 = 7776 - 1 = 7775$

Hence, option b) is correct.

Question 5

In a B⁺ tree indexing, size of the search key field is 10 bytes. If the block size is 2 kilobytes and the block pointer size is 14 bytes, calculate the order of a non-leaf node?

a) 204

1KB = 1024 bytes

b) 146

c) 85

d) 58

Answer: c)

Explanation: A B⁺ Tree of order m means every node has at most m children.

So, there will be only m number of block pointers and $(m - 1)$ number of key value in any non-leaf nodes.

P_1	K_1	P_2	\dots	P_{n-1}	K_{n-1}	P_n
-------	-------	-------	---------	-----------	-----------	-------

According to the question, maximum size of a non-leaf node ≤ 2 KB

size of search key field * $(m - 1)$ + block pointer * $m \leq 2048$

$10 * (m - 1) + 14 * m \leq 2048$

$10m - 10 + 14m \leq 2048$

$24m \leq 2058$

So, order of non-leaf node $(m) = 85$

Hence, option c) is correct.

Question 6

A relational table is saved in the disk and indexed with **bitmap index**. Size of the **bitmap indexed** file is 512 bytes. If there are 16 distinct values in the index column, how many rows are there in the relation table?

- a) 256
- b) 128
- c) 64
- d) 32

Answer: a)

Explanation: A **bitmap index** on a n row table with k distinct values in the index column will have $n \times k$ bits in the bitmap.

Therefore, $n \times k$ bits = The size of the **bitmap index** = 512 bytes

So, $n = (512 \times 8) \div 16 = 256$

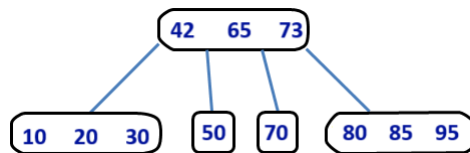
So, option (a) is correct.

Total size = n rows (bits) * k distinct values

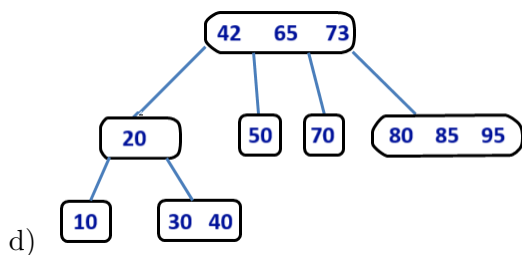
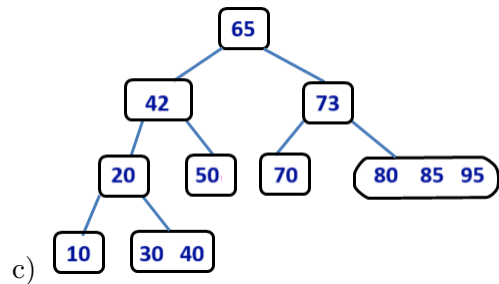
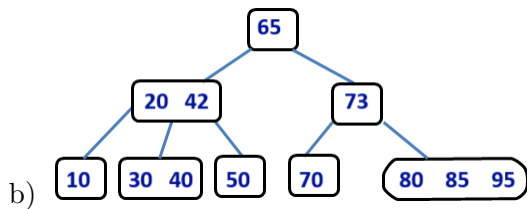
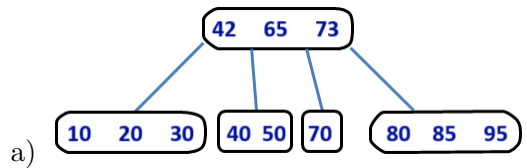
$$512 = (n/8) * 16$$

Question 7

Consider the following 2-3-4 Tree:

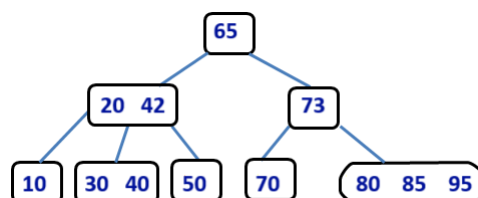


If we want to add a new key 40, what will be the resultant tree?



Answer: b)

Explanation: When we want to insert 40, it will come into the left-most sub-tree of 42. There are already 3 elements in that node. So, the node will split and 20 will go up into the root. In root there are already 3 elements, it also splits. Therefore, after adjustment of all pointers, the tree will look like:



Hence, option b) is correct.

Question 8

A hash table contains 10 (indices 0 to 9) buckets and uses linear probing to resolve collisions. The key values are integers and the hash function used is ($\text{key} \% 10$). Given the following input (221, 334, 447, 771, 574, 427, 673, 841), which of the following statements is true?

- a) 334, 841 hash to the same bucket.
- b) 221, 771, 841 are stored in buckets 1, 2, 3.
- c) 673, 334 are stored in buckets 3 and 4.
- d) 574 is stored in bucket 4.

Answer: c)

Explanation: After inserting all keys, the buckets will be:

	221	771	673	334	574	841	447	427	
0	1	2	3	4	5	6	7	8	9

Hence, option c) is correct.

Question 9

Consider the following table:

<u>ID</u>	PARTICIPANT	CATEGORY	GRADE
2322	AKSHAY NEOGI	DANCE	A
1221	JASMINE RAY	PAINTING	B
3242	ABHAY BERA	MUSIC	A
5850	SANTANU SEN	DANCE	A
1972	MITA SINH	MUSIC	B

The Bitmap indices of a CATEGORY and a GRADE are 00101 and 10110 respectively. Identify the value of CATEGORY and the GRADE.

- a) CATEGORY=MUSIC and GRADE=A
- b) CATEGORY=MUSIC and GRADE=B
- c) CATEGORY=DANCE and GRADE=B
- d) CATEGORY=DANCE and GRADE=A

Answer: a)

Explanation: CATEGORY MUSIC is present in 3rd and 5th tuple and absent in 1st, 2nd and 4th tuples. So, the bitmap indices for MUSIC is 00101. Similarly, GRADE A is present in 1st, 3rd and 4th tuples and is absent in 2nd and 5th tuples. So, the bitmap indices for GRADE A is 10110. Hence, option a) is correct.

Question 10

Suppose there is a relation `student(id, name, age)`, with a B^+ tree index with search key `id`. What is the worst-case cost of finding records satisfying $20 < id < 40$ using this index, in terms of the number of records retrieved m and the height of the tree is h ?

- a) $O(m \cdot h)$
- b) $O(2 \cdot h + m)$
- c) $O(\log h + m)$
- d) $O(h + m)$

Answer: d)

Explanation: B^+ Tree Structure

The B^+ tree index is based on the attribute `id`.

The height of the tree is h .

The tree will store the `id` values in sorted order, and the leaf nodes will contain pointers to the actual records.

Worst-Case Cost

To find records where `id` > 20 , a search in the B^+ tree will be required. This search will require traversing the tree from the root to a leaf node. In the worst case, this search requires $O(h)$ time, where h is the height of the tree.

Once we find the first record, we need to scan through the leaf nodes to retrieve all m number of records where the `id` value is < 40 , which requires $O(m)$ times.

Searching for all the records satisfying $20 < id < 40$: requires $O(h+m)$ time to traverse the B^+ tree. Hence, option d) is correct.



Course Name: Database Management System

Assignment 7 - Week 7 (Jan 2025)

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10

Total mark: $10 \times 2 = 20$

Question 1

Consider the following lock compatibility matrix where S denotes a shared lock and X denotes an exclusive lock:

	S	X
S	True	False
X	False	False

Which of the following statements about lock compatibility is (are) correct?

- a) A transaction holding an S lock on a data item allows other transactions to acquire an S lock but not an X lock on the same data item.
- b) A transaction holding an S lock on a data item allows other transactions to acquire an X lock on the same data item.
- c) A transaction holding an X lock on a data item allows other transactions to acquire another X lock on the same data item.
- d) A transaction holding an X lock on a data item prevents other transactions from acquiring an S lock on the same data item.

Answer: a), d)

Explanation: According to the lock compatibility matrix, a transaction holding an S lock allows other transactions to acquire S locks but does not permit X locks. Therefore, statement (b) and (c) are incorrect.

Question 2

Consider the following **schedule S** involving five transactions T_1 , T_2 , T_3 , T_4 , and T_5 :

T_1	T_2	T_3	T_4	T_5
	W(X)			
R(X)				
W(Y)				
		W(X)		
	R(Z)			
			W(Y)	
				R(X)
				W(Z)

R(X) denotes read operation on data item X by transaction T_i .

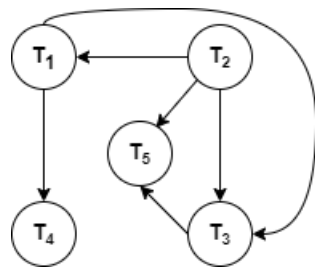
W(X) denotes write operation on data item X by transaction T_i .

Choose the correct option for the above **transaction schedule**.

- The **schedule** is neither **conflict serializable** nor **view serializable**.
- The **schedule** is both **conflict serializable** and **view serializable**.
- The **schedule** is only **view serializable**.
- The **schedule** is only **conflict serializable**.

Answer: (b)

Explanation: By drawing the **precedence graph** for the transactions and checking for cycles, we can determine the serializability of the schedule. If no cycles are found, the schedule is **conflict serializable**.



Additionally, **conflict serializable schedules** are always **view serializable**. So, option (b) is correct.

Question 3

Consider the following **schedule S** involving five transactions T_1 , T_2 , T_3 , T_4 and T_5 :

T_1	T_2	T_3	T_4	T_5
R(Z)				
R(X)				
		R(Y)		
	W(X)			
		W(X)		
			W(Z)	
				W(Z)

R(X) denotes read operation on data item X by transaction T_i .

W(X) denotes write operation on data item X by transaction T_i .

Identify the incorrect option(s) regarding the order of execution of all transactions in the above **schedule S**.

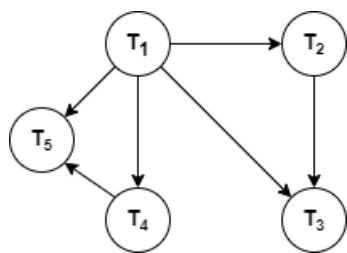
- a) $T_1 \rightarrow T_4 \rightarrow T_5 \rightarrow T_2 \rightarrow T_3$
- b) $T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_4 \rightarrow T_5$
- c) $T_1 \rightarrow T_2 \rightarrow T_4 \rightarrow T_5 \rightarrow T_3$
- d) $T_1 \rightarrow T_4 \rightarrow T_3 \rightarrow T_5 \rightarrow T_2$

Answer: d)

Explanation: If we draw the **precedence graph** of the transactions as shown in the following, we can observe that the graph has no cycle.

So, the above schedule is a **conflict serializable schedule**.

All **conflict serializable schedules** are **view serializable** too.



All possible **topological orderings** of the above **precedence graph** will be the possible **conflict serializable schedule**.

Hence, the incorrect order of execution of all transactions among the given options is: $T_1 \rightarrow T_4 \rightarrow T_3 \rightarrow T_5 \rightarrow T_2$

So, option (d) is the answer.

Question 4

Consider the following **schedule S** involving five transactions T_1 , T_2 , T_3 , T_4 and T_5 :

T_1	T_2	T_3	T_4	T_5
R(Z)				
	R(X)			
	W(X)			
		R(X)		
			R(Z)	
		W(X)		
			W(Z)	
W(Y)				
				W(Z)
	W(Y)			
				W(Y)

R(X) denotes read operation on data item X by transaction T_i .

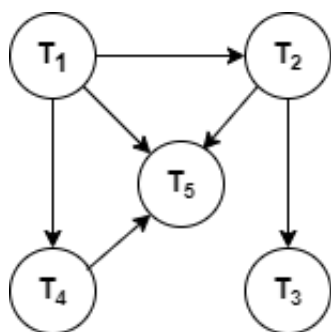
W(X) denotes write operation on data item X by transaction T_i .

Identify the possible number of **conflict serializable** schedules of the above **schedule S**.

- a) 1
- b) 2
- c) 3
- d) 5

Answer: d)

Explanation: If we draw the **precedence graph** of the **schedule**, we can observe that the graph has no cycle. Hence, the above schedule is **conflict serializable** schedules.



All possible **topological orderings** of the above **precedence graph** will be the possible **conflict serializable** schedule.

1. $T_1 \rightarrow T_4 \rightarrow T_2 \rightarrow T_5 \rightarrow T_3$
2. $T_1 \rightarrow T_4 \rightarrow T_2 \rightarrow T_3 \rightarrow T_5$
3. $T_1 \rightarrow T_2 \rightarrow T_4 \rightarrow T_3 \rightarrow T_5$
4. $T_1 \rightarrow T_2 \rightarrow T_4 \rightarrow T_5 \rightarrow T_3$
5. $T_1 \rightarrow T_2 \rightarrow T_3 \rightarrow T_4 \rightarrow T_5$

Hence, option d) is correct.

Question 5

Consider the following **schedule S** of transactions T_1 and T_2 .

The read operation on data item A is denoted by **read(A)** and the write operation on data item A is denoted by **write(A)**.

T_1	T_2
read(A)	
A:=A-500	
	read(C)
write(A)	
read(B)	
	temp:=C*0.5
	C:=C-temp
B:=B+500	
	write(C)
write(B)	
	read(B)
	B:=B+temp
	write(B)

Which of the following is **TRUE** about the **schedule S**?

- a) S is serializable both as T_1, T_2 and T_2, T_1 .
- b) S is not serializable neither as T_1, T_2 nor T_2, T_1 .
- c) S is serializable only as T_1, T_2 .
- d) S is serializable only as T_2, T_1 .

Answer: c)

Explanation: First, swap all non-conflicting instructions of the above schedule S.

T_1 and T_2 both are working on a same data item.

So, S is serializable only as T_1, T_2 .

Hence, option (c) is correct.

Question 6

Consider the following **schedule S**.

T1	T2	T3
		R(Y)
	W(Y)	
R(Y)		
		W(Y)
W(Y)		

R(Y) denotes read operation on data item Y by Transaction T_i .

W(Y) denotes write operation on data item Y by Transaction T_i .

Identify the possible number of **view serializable** schedule of the above schedule S.

- a) 1
- b) 2
- c) 4
- d) 6

Answer: a)

Explanation: Step 1: Final Update on data item: This is done on Y by T1.

Since the final update on Y is made by T1, the transaction T1 must execute after the transactions T2 and T3 So, $(T2, T3) \rightarrow T1$.

Step 2: Initial Read + Which transaction updates after read?

Transaction T3 does the initial read Y; then, T2 updates on it.

Hence, the dependency is: $T3 \rightarrow T2$

Step 3: Write Read Sequence: T2 writes Y and then T1 reads Y.

Hence, the dependency is as follows. $T2 \rightarrow T1$

Thus, the only possible way, this is possible is: $T3 \rightarrow T2 \rightarrow T1$.

Hence, total possible **view serializable** schedule of S= 1.

So, option (a) is correct.

Question 7

Consider two transactions given below where `lock-X(A)` denotes T_i has obtained an **Exclusive-mode** lock on item A and `lock-S(A)` denotes T_i has obtained a **Shared-mode** lock on item A.

T_1	T_2
<code>lock-X(A)</code>	<code>lock-X(A)</code>
<code>read(A)</code>	<code>read(A)</code>
<code>lock-X(B)</code>	<code>A:= A-100</code>
<code>read(B)</code>	<code>write(A)</code>
<code>B:= B+100</code>	<code>lock-S(B)</code>
<code>write(B)</code>	<code>read(B)</code>
<code>lock-S(C)</code>	<code>lock-S(C)</code>
<code>read(C)</code>	<code>read(C)</code>
<code>unlock(C)</code>	<code>unlock(B)</code>
<code>commit</code>	<code>unlock(C)</code>
<code>unlock(A)</code>	<code>commit</code>
<code>unlock(B)</code>	<code>unlock(A)</code>

Which of the following statement is (are) true?

- a) T_2 follows the rigorous two-phase locking protocol, but T_1 follows the strict two-phase locking protocol only .
- b) T_1 follows the rigorous two-phase locking protocol, but T_2 follows the strict two-phase locking protocol only.
- c) Both T_1 and T_2 follow the strict two-phase locking protocol.
- d) Both T_1 and T_2 follow the rigorous two-phase locking protocol.

Answer: c)

Explanation: Transaction T_1 first unlocks all **Shared-mode** lock and then commits. After commit, it unlocks **Exclusive-mode** lock . That is why, T_1 follows the **strict two-phase locking protocol**. Transaction T_2 first unlocks all **Shared-mode** lock and then commits. After commit, it unlocks **Exclusive-mode** lock . That is why, it follows the **strict two phase locking protocol** only. So, option (c) is correct.

Question 8

Consider the following schedule S.

T_1	T_2
R(X)	
W(X)	
R(Y)	
W(Y)	
COMMIT	
	R(X)
	W(X)
	R(Y)
	W(Y)
	COMMIT

R(X) denotes read operation on data item X by Transaction T_i .

W(X) denotes write operation on data item X by Transaction T_i .

Choose the correct options for the above schedule.

- a) The schedule is only **recoverable schedule**.
- b) The schedule is only **cascadeless schedule**.
- c) The schedule is **recoverable schedule** and **cascadeless schedule** both.
- d) The schedule is neither **recoverable nor cascadeless schedule**.

Answer: c)

Explanation: Recoverable schedule: If a transaction T_j reads a data item previously written by a transaction T_i , the commit operation of T_i must appear before the commit operation of T_j .

Cascadeless schedules: For each pair of transactions T_i and T_j such that T_j reads a data item previously written by T_i , the commit operation of T_i appears before the read operation of T_j .

Here, T_2 read all data items after commit operation of T_1 . Hence, the schedule is recoverable as well as cascadeless.

So, option (c) is correct.

Question 9

Consider the following schedule S.

T1	T2	T3
R(X)		
W(X)		
	R(X)	
	W(X)	
	R(Y)	
	W(Y)	
		R(Y)
		W(Y)
	abort	

R(X) denotes a read operation on data item X by transaction T_i .

W(X) denotes a write operation on data item X by transaction T_i .

Transaction T3 commits before T2 aborts.

Identify the correct statement(s) based on the above schedule S.

- a) If T2 fails (aborts), only T1 will be rolled back, while T3 will remain unaffected.
- b) If T2 fails (aborts), only T3 will be rolled back, while T1 will remain unaffected.
- c) If T2 fails (aborts), no other transaction will be rolled back.
- d) If T2 fails (aborted), both transactions T1, and T3 must also be rolled back.

Answer: d)

Explanation:

- In the given schedule, T3 depends on the data item Y written by T2. If T2 aborts, T3 must also roll back to ensure consistency.
- In the given schedule, T2 depends on the data item X written by T1. If T2 aborts, T1 must also roll back to ensure consistency.

Thus, the correct answer is **(d)**: If T2 fails, both transactions T1, and T3 must also be rolled back.

Question 10

Suppose in a database, there are three transactions T_1 , T_2 , and T_3 with timestamps 20, 21, and 22 respectively. T_2 is holding some data items which T_1 and T_3 are requesting to acquire. Which of the following statement(s) is (are) correct in respect of **Wait-Die Deadlock Prevention Scheme**?

- a) Transaction T_1 will rollback.
- b) Transaction T_3 will wait for T_2 to release the data item.
- c) Transaction T_1 will wait for T_2 to release the data item and Transaction T_3 will rollback.
- d) Both Transactions T_1 and T_3 will rollback.

Answer: c)

Explanation: In **Wait-Die** Deadlock Prevention scheme:

Older transaction may **wait** for younger one to release data item. (older means smaller timestamp)

Younger transactions never wait for older ones; they are **rolled back** instead.

Transaction T_1 is older than T_2 , hence T_1 will wait until the data item is released. Transaction T_3 is younger than Transaction T_2 , so it will be rolled back.

Hence, option c) is correct.



Course Name: Database Management System

Assignment 8 - Week 8 (Jan 2025)

TYPE OF QUESTION: MCQ/MSQ

Number of questions: 10

Total mark: $10 \times 2 = 20$

Question 1

Identify the cost estimation of a query evaluation plan, if 9000 blocks are required to be transferred from the disk and the required number of disk seeks are 25.

- Time to transfer one block: $t_T = 5$ milliseconds.
- Time for one seek: $t_S = 0.4$ seconds.

- a) 40 Seconds
- b) 45 Seconds
- c) 50 Seconds
- d) 55 Seconds

Answer: d)

Explanation: Cost for b block transfers plus S seeks will be $(b * t_T + S * t_S)$ seconds
 $= (9000 * 5 * 10^{-3}) + (25 * 0.4)$ seconds
 $= (45 + 10)$ Seconds
 $= 55$ Seconds

For more details refer to 38.12 of lecture material.

Hence, option d) is the answer.

Question 2

Assume an immediate database modification scheme. Consider the following log records for transactions T0, T1, T2, T3 and T4:

steps	Details of log
1	$\langle T0, start \rangle$
2	$\langle T0, A, 500, 600 \rangle$
3	$\langle T1, start \rangle$
4	$\langle T1, B, 300, 500 \rangle$
5	$\langle T1, commit \rangle$
6	$\langle T2, start \rangle$
7	$\langle checkpoint\{T0, T2\} \rangle$
8	$\langle T3, start \rangle$
9	$\langle T2, C, 200, 400 \rangle$
10	$\langle T3, D, 700, 900 \rangle$
11	$\langle T2, commit \rangle$
12	$\langle T3, commit \rangle$
13	$\langle T4, start \rangle$
14	$\langle T4, E, 300, 700 \rangle$

If there is a crash just after step 14 and the recovery of the system is successfully completed, identify the **correct** action for the above scenario.

- a) After recovery completion, value of A will be 600.
- b) After recovery completion, value of C will be 200.
- c) After recovery completion, value of D will be 900.
- d) After recovery completion, value of E will be 300.

Answer: c), d)

Explanation: In the immediate database modification scheme, during recovery after a crash, a transaction needs to be redone if and only if both $\langle T_i, start \rangle$, $\langle T_i, commit \rangle$ are present in the log. otherwise undo is required.

Any transactions that committed before the last checkpoint should be ignored (updates already output to disk due to the checkpoint).

Redo list contains transaction $\{T2, T3\}$ and undo list contain transactions $\{T0, T4\}$ and for transaction $\{T1\}$ no need any action because it is already committed before checkpoint.

As per the process of transaction recovery, options (c) and (d) are correct.

Question 3

Let us consider the following statistics for two relations `Instructor` and `Job_Assignments`:

- Number of records of `Instructor`: $n_{\text{Instructor}} = 5050$.
- Number of blocks of `Instructor`: $b_{\text{Instructor}} = 30$.
- Number of records of `Job_Assignments`: $n_{\text{Job_Assignments}} = 1050$.
- Number of blocks of `Job_Assignments`: $b_{\text{Job_Assignments}} = 10$.

Let us consider a natural join of `Instructor` and `Job_Assignments` relations (`Instructor` \bowtie `Job_Assignments`). Identify the required number of **block transfers** in the worst case (enough memory only to hold one block of each relation) using **Nested-loop join** and assume `Instructor` as the outer relation.

- a) 40000 block transfers
- b) 40030 block transfer
- c) 50030 block transfers
- d) 50530 block transfers

Answer: d)

Explanation: Number of block transfers will be: $5050 \times 10 + 30 = 50530$, if `Instructor` is taken as the outer relation.

For more details refer to 38.31 of lecture material.

Question 4

Consider the following relational schema:

INSTRUCTOR(InstructorID, Name, HireDate)

COURSE(CourseID, CourseName, Credits)

JOB_ASSIGNMENTS(InstructorID, CourseID, JobTitle, StartDate, EndDate)

Two query trees are given below.

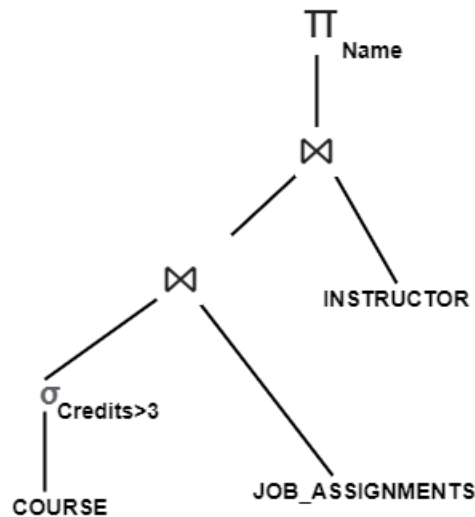


Figure 1:

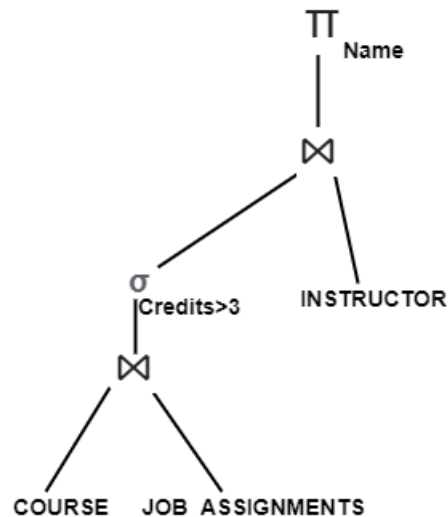


Figure 2:

Identify the correct statement for the above two query trees.

- a) Two query trees are equivalent, and the query tree of Figure 2 will lead to more efficient query processing.
- b) Two query trees are equivalent, and the query tree of Figure 1 will lead to more efficient query processing.
- c) Two query trees are equivalent, as identical operations (irrespective of their positions) are used in both trees.
- d) Two query trees are not equivalent as selection or projection operation cannot be carried out before or after the natural join operation.

Answer: b)

Explanation: Two query trees are equivalent, and Figure 1 will lead to more efficient query processing because performing the selection operation as early as possible reduces the size of the relation to be joined.

Hence, option (b) is correct.

Question 5

Assume deferred database modification scheme. Consider the following log records for transactions T1, T2, T3 and T4:

steps	Details of log
1	$\langle T1, start \rangle$
2	$\langle T1, A, 500, 800 \rangle$
3	$\langle T2, start \rangle$
4	$\langle checkpoint\{T1, T2\} \rangle$
5	$\langle T1, commit \rangle$
6	$\langle T3, start \rangle$
7	$\langle T2, B, 200, 400 \rangle$
8	$\langle T2, commit \rangle$
9	$\langle T4, start \rangle$
10	$\langle T3, C, 700, 300 \rangle$

If there is a crash just after step 10 and the recovery of the system is successfully completed, identify the **correct** action(s) for the above scenario.

- a) After recovery completion, the value of B is 400.
- b) After recovery completion, the value of C is 300.
- c) Redo list contain transactions $\{T1, T2\}$ and undo list contains $\{T3, T4\}$.
- d) Redo list contain transactions $\{T1, T2\}$ and undo list contains $\{ \}$.

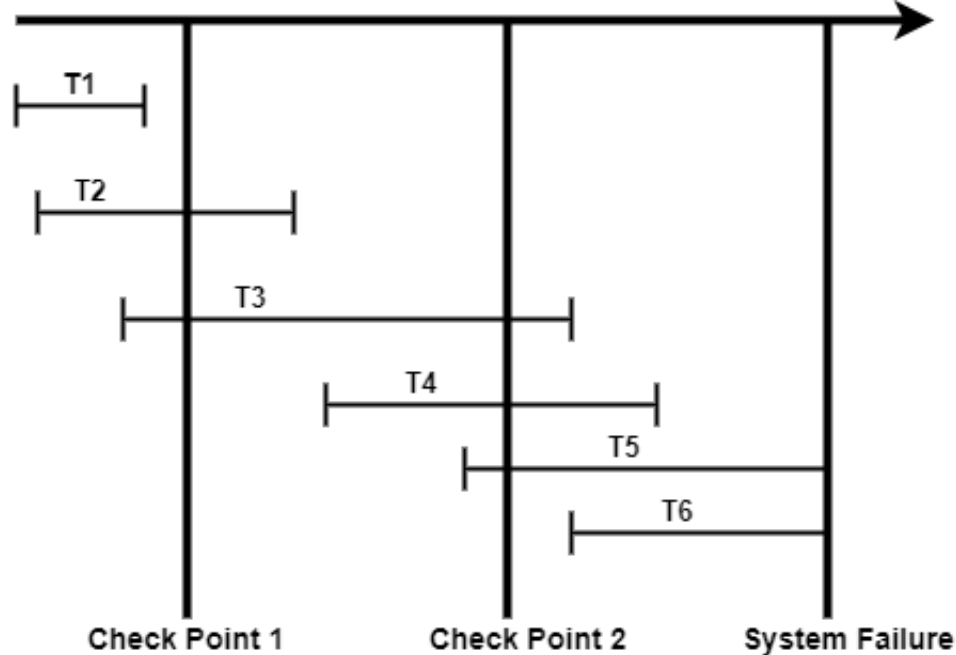
Answer: a), d)

Explanation: In the deferred database modification scheme, updates are not written to the database until the transaction commits. If a transaction commits, its updates are redone from the log. If a transaction does not commit, it is ignored (no undo is needed, since no changes were written to the database).

As per the process of transaction recovery, options a) and d) are correct.

Question 6

Consider the following state of transactions and the statements below.



1. T_1 , T_2 and T_3 can be ignored.
2. T_5 and T_6 need to be undone.
3. T_1 and T_2 can be ignored.
4. T_3 , T_4 and T_5 need to be redone.
5. T_3 and T_4 need to be redone.

Identify the correct group of statements from the options below.

- a) 1), 2), 3), 5)
- b) 1), 3), 4), 5)
- c) 1), 4), 5)
- d) 2), 3), 5)

Answer: d)

Explanation: Any transaction that is committed before the last checkpoint should be ignored. Therefore, T_1 , and T_2 can be ignored (updates already output to disk due to the last checkpoint).

Any transaction that is committed since the last checkpoint, needs to be redone. Hence, T_3 and T_4 are to be redone.

Any transaction that was running at the time of failure, needs to be undone and restarted. Hence, T_5 and T_6 are to be undone.

Hence, option (d) is correct.

Question 7

Consider the following relational schema:

INSTRUCTOR(InstructorID, Name, HireDate)

COURSE(CourseID, CourseName, Credits)

JOB_ASSIGNMENTS(InstructorID, CourseID, JobTitle, StartDate, EndDate)

Four relational algebra queries are given below:

Q1: $\sigma_{\text{StartDate}='2023-01-01'}(\text{INSTRUCTOR} \bowtie \text{JOB_ASSIGNMENTS})$

Q2: $\pi_{\text{StartDate}, \text{Name}}(\text{INSTRUCTOR} \bowtie \text{JOB_ASSIGNMENTS})$

Q3: $\pi_{\text{Name}}(\text{INSTRUCTOR} \bowtie \text{JOB_ASSIGNMENTS})$

Q4: $\pi_{\text{Name}}(\sigma_{\text{INSTRUCTOR.InstructorID}=\text{JOB_ASSIGNMENTS.InstructorID}}(\text{INSTRUCTOR} \times \text{JOB_ASSIGNMENTS}))$

Identify the correct options from the options given below.

- a) Q1 is equivalent to Q2.
- b) Q1 is not equivalent to Q2.
- c) Q3 is equivalent to Q4.
- d) Q3 is not equivalent to Q4.

Answer: b), c)

Explanation: Q1 and Q2 will not give the same result because Q2 will project only the StartDate and Name columns over the natural join of INSTRUCTOR and JOB_ASSIGNMENTS, whereas Q1 selects records where StartDate is '2023-01-01'.

Q3 and Q4 will give the same result because $\sigma_{\text{INSTRUCTOR.InstructorID}=\text{JOB_ASSIGNMENTS.InstructorID}}$ followed by $(\text{INSTRUCTOR} \times \text{JOB_ASSIGNMENTS})$ is equivalent to $(\text{INSTRUCTOR} \bowtie \text{JOB_ASSIGNMENTS})$.

Hence, options b) and c) are correct.

Question 8

Consider the following relational schema:

INSTRUCTOR(InstructorID, Name, HireDate)

COURSE(CourseID, CourseName, Credits)

JOB_ASSIGNMENTS(InstructorID, CourseID, JobTitle, StartDate, EndDate)

A relational algebra expression is given below:

$\Pi_{\text{Name}}(\Pi_{\text{InstructorID}, \text{Name}}(\sigma_{\text{Credits}=3}(\text{INSTRUCTOR} \bowtie \text{JOB_ASSIGNMENTS} \bowtie \text{COURSE})))$

Identify the most optimized relational algebra expression equivalent to the above relational algebra expression.

- a) $\Pi_{\text{Name}, \text{InstructorID}}(\sigma_{\text{Credits}=3}(\text{INSTRUCTOR} \bowtie \text{JOB_ASSIGNMENTS} \bowtie \text{COURSE}))$
- b) $\Pi_{\text{Name}, \text{InstructorID}}(\text{INSTRUCTOR} \bowtie (\text{JOB_ASSIGNMENTS} \bowtie (\sigma_{\text{Credits}=3}(\text{COURSE}))))$
- c) $\Pi_{\text{Name}, \text{InstructorID}}(\text{INSTRUCTOR} \bowtie (\text{JOB_ASSIGNMENTS} \bowtie \Pi_{\text{CourseID}}(\sigma_{\text{Credits}=3}(\text{COURSE}))))$
- d) $\Pi_{\text{Name}}(\text{INSTRUCTOR} \bowtie (\text{JOB_ASSIGNMENTS} \bowtie (\sigma_{\text{Credits}=3}(\text{COURSE}))))$

Answer: d)

Explanation: According to the transformation rules, only the last in a sequence of projection operations is needed, so the others can be omitted.

It is optimal to perform the selection operation as early as possible to reduce the size of the relation to be joined.

Hence, option (d) is correct.

Question 9

Consider the following two relational algebra expressions (RA) given below:

RA I: $\Pi_{A,B}(P \cup Q) = \Pi_{A,B}(P) \cup \Pi_{A,B}(Q)$

RA II: $((P \bowtie Q) \bowtie R) = (P \bowtie (Q \bowtie R))$

where P, Q, and R are relational algebra expressions.

Identify the correct statement(s) from the following.

- a) Both RA I and RA II are true.
- b) Both RA I and RA II are false.
- c) RA I is true but RA II is false.
- d) RA I is false but RA II is true.

Answer: a)

Explanation: RA I follows the projection distribution property over union: projecting after a union is equivalent to taking projections first and then performing a union. Thus, RA I is valid.

RA II is correct because natural join operations are associative.

Thus, option a) is correct.

Question 10

Consider the log record of Transaction T1 with two operation instances O1 and O2 used in recovery system with early lock release, B+ tree based concurrency control.

Step	Operation
1	$\langle T1, \text{start} \rangle$
2	$\langle T1, X, 200, 400 \rangle$
3	$\langle T1, O1, \text{operation-begin} \rangle$
4	$\langle T1, Y, 100, 500 \rangle$
5	$\langle T1, O1, \text{operation-end}, (Y, -400) \rangle$
6	$\langle T1, O2, \text{operation-begin} \rangle$
7	$\langle T1, Z, 500, 800 \rangle$
8	crash or abort here

Choose the correct set of log entries for the recovery of transactions.

a)	$\langle T1, Z, 500 \rangle$
	$\langle T1, Y, 500, 100 \rangle$
	$\langle T1, O1, \text{operation-abort} \rangle$
	$\langle T1, X, 400 \rangle$
	$\langle T1, \text{abort} \rangle$

b)	$\langle T1, Z, 500 \rangle$
	$\langle T1, Y, 500, 100 \rangle$
	$\langle T1, O1, \text{operation-abort} \rangle$
	$\langle T1, X, 200 \rangle$
	$\langle T1, \text{abort} \rangle$

c)	$\langle T1, Z, 500 \rangle$
	$\langle T1, Y, 100, 500 \rangle$
	$\langle T1, O1, \text{operation-abort} \rangle$
	$\langle T1, X, 200 \rangle$
	$\langle T1, \text{abort} \rangle$

d)	$\langle T1, Z, 800 \rangle$
	$\langle T1, Y, 500, 100 \rangle$
	$\langle T1, O1, \text{operation-abort} \rangle$
	$\langle T1, X, 400 \rangle$
	$\langle T1, \text{abort} \rangle$

Answer: b)

Explanation: Step i: Scan the log records backward.

Step ii: For the step 7 update of Z was part of O2, undone physically during recovery since O2 did not complete; so, physical undo is required for operation O2 on the variable Z. That means, we have to add log $\langle T2, Z, 500 \rangle$. Step iii: For the step 5 $\langle T1, O1, \text{operation-end}, (Y, -400) \rangle$ is found; so, logical undo is required for operation O1 on the variable Y using the information (subtraction of 400). That means, we have to delete the previous modifications on Y. So, add the following logs for steps 5, 4 and 3 respectively: $\langle T1, Y, 500, 100 \rangle$ and $\langle T1, O1, \text{operation-abort} \rangle$.

Step iv: For step 2, add the log $\langle T1, X, 200 \rangle$.

Step v: For step 4, add the log $\langle T1, \text{abort} \rangle$.

Hence, option (b) is correct.