GATE CS Topic wise Questions

Database

YEAR 2000

Question. 1

Given the relations

employee (name, salary, deptno), and

department (deptno, deptname, address)

Which of the following queries cannot be expressed using the basic relational algebra operations ($\sigma, \pi, \bowtie, \cup, \cap, -$) ?

(A) Department address of every employee

(B) Employee whose name is the same as their department name

(C) The sum of all employee salaries

(D) All employees of a given department

SOLUTION

Question. 2

Given the following relation instance.

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

Which of the following functional dependencies are satisfied by the instance ?

(A) $XY \rightarrow Z$ and $Z \rightarrow Y$ (B) $YZ \rightarrow X$ and $Y \rightarrow Z$

(C) $YZ \rightarrow X$ and $X \rightarrow Z$ (D) $XZ \rightarrow Y$ and $Y \rightarrow X$

SOLUTION

Question. 3

Given relations $r(w,x)$ and $s(y,z)$, the result of select distinct w,x from r,s:

is guaranteed to be same as r, provided :

(A) r has no duplicates and s is non empty

(B) r and s have no duplicates

(C) s has no duplicates and r is non empty

(D) r and s have the same number of tuples

YEAR 2001

Question. 4

Consider a schema R(A,B,C,D) and functional dependencies $A \rightarrow B$ and $C \rightarrow D$. Then the decomposition of R into$R_1(AB)$ and $R_2(CD)$ is :

(A) Dependency preserving and lossless join

(B) Lossless join but not dependency preserving

(C) Dependency preserving but not lossless join

(D) Not dependency preserving and not lossless join

SOLUTION

$R(A,B,C,D)$

$R_1(AB) \Rightarrow A \rightarrow B$, $C \rightarrow D$

$R_1(AB)$ & $R2(CD)$.

So dependencies are preserved during decomposition.

But for lossless join.

$R_1(AB) \cap R_2(CD) \rightarrow R_1(AB)$

$R_1(AB) \cap R_2(CD) \rightarrow R_2(CD)$
Since $A \rightarrow B \ \& C \rightarrow D$
So lossless join not possible.
Hence (C) is correct option.

Question. 5

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

(A) List all vertices adjacent to a given vertex.

(B) List all vertices which have self loops

(C) List all vertices which belong to cycles of less than three vertices

(D) List all vertices reachable from a given vertex

SOLUTION

The database contains the adjacency list of the graph. So relation algebra with face problems when while calculating the length of cycle self loops come, then the query would execute in one tuple only.
Hence (C) is correct option.

Question. 6

Let r and s be two relations over the relation schemes R and S respectively, and let A be an attribute in R. Then the relational algebra expression $\sigma_{A=a}(r \bowtie s)$ is always equal to :

(A) $\sigma_{A=a}(r)$ (B) r

(C) $\sigma_{A=a}(r \bowtie s)$ (D) None of the above

SOLUTION

Given query performs natural join between r & s, & then project attribute A where $A = a$.
This same result is produced by the query $\sigma_{A=a}(r) \bowtie s$.
This query selects attribute $A = a$ from r & then performs join operation results are same.
Hence (C) is correct option.

Question. 7

R,(A,B,C,D) is a relation. Which of the following does not have a

lossless join, dependency preserving BCNF decomposition ?

(A) $A \rightarrow B, B \rightarrow CD$ (B) $A \rightarrow B, B \rightarrow C, C \rightarrow D$

(C) $AB \rightarrow C, C \rightarrow AD$ (D) $A \rightarrow BCD$

SOLUTION

$R(A, B, C, D)$

In option (A) & (B) there exists transitive functional dependency, so incorrect for $BCNF$.

In option (C)

$AB \rightarrow C \quad C \rightarrow AD$

So $AB \rightarrow AD$

$B \rightarrow D$

In option (D) all attributes depend upon same key so in $BCNF$.

Hence (D) is correct option.

Question. 8

Which of the following relational calculus expressions is not safe ?

(A) $\{r \mid \exists u \in R_1(t[A]) = u[A] \hat{} \neg \exists s \in R_2(t[A] = s[A])\}$

(B) $\{r \mid \forall u \in R_1(u[A]) = "x" \Rightarrow \exists s \in R_2(t[A] = s[A] \hat{} \ s[A] = u[A]))\}$

(C) $\{t \mid \neg(t \in R_1)\}$

(D) $\{t \mid \exists u \in R_1(t[A] = u[A]) \hat{} \ \exists s \in R_2 9t[A] = s[A])\}$

SOLUTION

In option (C) $\{t\Gamma(t \in R_1)\}$ is not safe since no criteria for selection of tuple has been given.

Hence (C) is correct option.

Question. 9

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

Create table gaq

(Ib integer not null

ub integer not null

primary key Ib

foreign key (ub) references geq on delete cascade):

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

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

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

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

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

SOLUTION

Tuple (x, y) is deleted.
Here $y \geq x$
lb is primary key
ub is foreign key

Table A (geq)

X	Y
x	y

→ deleted

Table B (geq)

X	Y
z	w

Since y refer to same key z. & in table A x a primary key is deleted so all enteries from table B will also be deleted.
And also all enteries referencing it in y will also be deleted.
So a tuple (z, w) with $w < x$ will also be deleted.
Hence (C) is correct option.

YEAR 2002

Question. 10

Relation R with an associated set of functional dependencies, F, is decomposed into $BCNF$. The redundancy (arising out of functional dependencies) in the resulting set of relations is.

(A) Zero

(B) More than zero but less than that of an equivalent $3NF$ decomposition

(C) Proportional to the size of F^+

(D) Indetermine.

SOLUTION

BCNF (Boyce Cold Normal Form) Since R is decomposed in *BCNF* then no redundancy in the resulting set may occur and all transitive dependencies are removed.
Hence (A) is correct option.

Question. 11

With regard to the expressive power of the formal relational query languages, which of the following statements is true ?

(A) Relational algebra is more powerful than relational calculus.

(B) Relational algebra has the same power as relational calculus.

(C) Relational algebra has the same power as safe relational calculus.

(D) None of the above.

SOLUTION

Expressive power is the capacity of formal query languages to express various query statements, so relational algebra is as powerful as relational calculus only if calculus is safe relational calculus.
Hence (C) is correct option.

Question. 12

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

(A) 16 (B) 42

(C) 43 (D) 44

SOLUTION

$$\text{Size of 1 record of index} = 8 + 4 = 12 \text{ bytes.}$$
$$\text{Let no. of pointers required} = P$$
$$\text{No. of index values per block} = P - 1$$
$$\text{So} \quad (P-1)8 + 4P = 512$$
$$12P = 520$$
$$P \cong 44$$

Hence (D) is correct option.

Question. 13

Relation R is decomposed using a set of functional dependencies, F,and relation S is decomposed using another set of functional dependencies, G. One decomposition is definitely $BCNF$, the other is definitely. $3NF$, but it is not known which is which. To make a guaranteed identification, which one of the following tests should be used on the decompositions ? (Assume that the closures of F and G are available).

(A) Dependency-preservation (B) Lossless-join

(C) $BCNF$ definition (D) $3NF$ definition

SOLUTION

$BCNF$ is more stricter than $3NF$. So if we check database for $BCNF$ then it will definitely be in $3NF$.
Hence (C) is correct option.

Question. 14

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

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

(A) A functionally determines B and B functionally determines C

(B) A functionally determines B and B does not functionally determines C.

(C) B does not functionally determines C

(D) A does not functionally B and B does not functionally determines.

SOLUTION

A is said to Be functionally determined by B only is A has same

value for same values of B for every corresponding tuple.

In given database.

A	B
1	1
1	1
2	3
2	3

here if A is 1 then B is necessarily 1 & if A is 2 then B is 3.

So B is functionally determined by A.

But this relationship is not true for any other pair.

Hence (B) is correct option.

YEAR 2003

Question. 15

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

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

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

(C) A transaction read a data item after it is written by an committed transaction

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

SOLUTION

(A) read & then write.

(B) read after read

(C) read after committed write

(D) read after uncommitted write

Option (D) is a condition for inconsisting

Option (D) is correct option

Question. 16

Consider the following SQL query

select distinct $a_1, a_2, \ldots\ldots a_n$

from $r_1, r_2, \ldots\ldots r_m$

where P

For an arbitrary predicate P,this query is equivalent to which of the following relational algebra expressions?

(A) $\underset{a_1, a_2, \ldots a_n}{\Pi} \sigma_\rho (r_1 \times r_2 \times \ldots\ldots r_m)$

(B) $\underset{a_1, a_2, \ldots a_n}{\Pi} \sigma_\rho (r_1 \rhd\lhd r_2 \rhd\lhd \ldots\ldots \rhd\lhd r_m)$

(C) $\underset{a_1, a_2, \ldots a_n}{\Pi} \sigma_\rho (r_1 \cup r_2 \cup \ldots\ldots \cup r_m)$

(D) $\underset{a_1, a_2, \ldots a_n}{\Pi} \sigma_\rho (r_1 \cap r_2 \cup \ldots\ldots \cap r_m)$

SOLUTION

Here the *SQL* query selects distinct attributes, in relation algebra it doesn't require any symbol but only projection. P is the predicate in relation algebra, "from" has multiple relations, this is join in relation algebra (RA).

$\underset{a_1, a_2 \ldots\ldots a_n}{\Pi} \sigma_\rho (r_1 \times r_2 \times \ldots\ldots \times r_m)$

Hence (A) is correct option.

Question. 17

Consider the following functional dependencise in a database:

Data_of_Birth→Age Age→Eligibility

Name→Roll_number Roll_number→Name

Course_number→Course_name Course_number→Instructor

(Roll_number,Course_number)→Grade

The relation(Roll)number,Name,Date_of_brith,Age)is

(A) in second normal normal form but not in third normal form

(B) in third normal form but not in BCNF

(C) in BCNF

(D) in none of the above

SOLUTION

Date of Birth → Age

Age → Eligibility so transitively

Date of Birth → Eligibility

Name ⇄ Roll no.

Given relation ⇉ (Roll no. Name, Date of Birth, Age) Roll no. & Name are keys.

1. Age is functionally dependent upon date of birth which is non key, so $2NF$ is not applicable.
2. In $3NF$, every non key should be non transitively or fully dependent upon all the keys, but here this is not the case, so not in $3NF$.
3. In $BCNF$ every non trivial functional dependency in the relation is a dependency on a super key. So also not in $BCNF$.

So it is in INF only, since no repeating tuples.

Hence (D) is correct option

Question. 4

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

Students:(Roll_number,Name,Date_of_birth)

Courses:(Course_number,Course_name,Instructor)

Grades:(Roll_number,Course_number,Grade)

select distrinct Name

from Students, Courses, Grades

Where Students,Roll_number=Grades. Toll_number

and Courses. Instructor=Korth

and Courses. Course_number=Grades. Course_number

and Grades.grade=A

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

(A) Names of students who have got an A grade in all courses taught by Korth

(B) Names of students who have got an A grade in all courses

(C) Name of students who have got an A grade in at least one of the courses taught by Korth

(D) None of the above

SOLUTION

Names are unique since selected distinct, so one person will be selected only once, inspite of the has grade A in any no. of courses taught by korth.
So if a name appearing, it is not necessary that he had grade A in all the courses, but it has A in at least 1 course.
Hence (C) is correct option.

Question. 5

Consider three data items $D1$,$D2$ and $D3$ and the following execution schedule of transactions $T1, T2$ and $T3$. In the diagram, $R(D)$ and $W(D)$ denote the actions reading and writing the data item D respectively.

T1	T2	T3
	R(D3);	
	R(D2);	
	R(D2);	
		R(D2);
		R(D3);
R(D1);		
R(D1);		
		W(D2);
		W(D3);
	R(D1);	
R(D2);		
W(D2);		
	W(D1);	

(A) The schedule is serializable as T2; T3;T1;

(B) The schedule is serializable as T2; T1;T3;

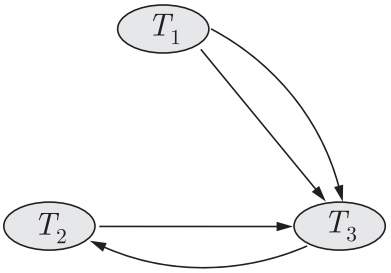
(C) The schedule is serializable as T3; T2; T1;

(D) The schedule is not serializable

SOLUTION

Let us draw a flow diagram between the three schedules, an edge between two schedules is there if there exist read-write or write-write

dependency in between them.
So we draw the graph.



Since there exist a cycle between $T1$ & $T3$ so there is no serializability.
Hence (D) is correct option

YEAR 2004

Question. 6

Let $R_1(A, B, C)$ and $R_2(D, E)$ be two relation schema, where the primary keys are shown underlined, and let C be a foreign key in R_1 referring to R_2. Suppose there is no violation of the above referential integrity constraint in the corresponding relation instances $r1$ and $rl2$. Which one of the following relational algebra expressions would necessarily produce an empty relation?

(A) $\Pi_D(r_1) - \Pi_C(r_1)$ (B) $\Pi_C(r_1) - \Pi_D(r_1)$

(C) $\Pi_D(r_1 \bowtie_{C \neq D} R_2) - \Pi_C(r_1)$ (D) $\Pi_C(r_1 \bowtie_{C = D} R_2)$

SOLUTION

C is an attribute in R_1 but D is a key in K_2.
So consider $\Pi_C(r_1) - \Pi_D(r_2)$
So the result of this query would be all those tuples which are in $\Pi_C(r_1)$ but not in $\Pi_D(r_2)$.
Since D is a key so it has all the possible values of C.
So difference would always be empty.
Hence (B) is correct.

Question. 7

Consider the following relation schema pertaining to a students database:

Student(rollno,name,address)

Enroll(rollno,courseno, coursename)

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

(A) 8,8 (B) 120,8

(C) 960,8 (D) 960,120

SOLUTION

The boundary cases are when either all the tuples of Enroll table belong to one roll no. , so there can be at most 8 roll no & courses no. combinations or the other case is when all the tuples belong to different roll no. this also has 8 tuples.

So (8,8) = (max,min)

Hence (A) is correct option

Question. 8

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

(i) getId is implemented in the superclass

(ii) getId is implemented in the suclass

(iii) getName is an abstract function in the superclass

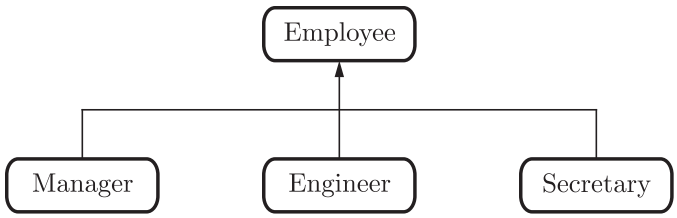
(iv) getName is implemented in the superclass

(v) getName is implemented in the subclass

(vi) getSalary is an abstract function in the superclass

(vii) getSalary is implemented in the superclass

(viii) getSalary is implemented in the subclass



Choose the best design

(A) (i),(iv),(vi).(viii)
(B) (i),(iv),(vii)
(C) (i),(iii),(v),(vi),(viii)
(D) (ii),(v),(viii)

SOLUTION

Employee *ID* & employee name are independent of subclasses; manager engineer or secretary, so these should be implemented in super class, but salary is part of employee dependent, so abstract get salary should be in super class & its actual implementation in subclass.

Hence (A) is correct option.

Question. 9

The relation scheme student Performance (name, courselNo, rollNo, grade) has the following functional dependencies:

name, courseNo →grade
RollNo, courseNo →grade
name →rollNo
rollNo →name

The highest normal form of this relation scheme is

(A) 2 NF
(B) 3NF
(C) BCNF
(D) 4 NF

SOLUTION

Name, course no. → grade
Roll no. course no. → grade

Name, Roll no are candidate keys grade depend upon both super keys.

So it is $2NF$, non key attribute grade depend upon super keys fully.

But not in $3NF$, since grade is not fully dependent upon all candidate keys.

Hence (A) is correct option.

Question. 10

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

$$\Pi_{\text{name}}(r_{\text{sex}=\text{females}}(\text{Student}))\, P_{\text{name}}(\text{Student} \underset{\substack{(\text{sex}=\text{female} \\ \wedge\ x=\text{male} \\ \wedge \text{marks} \le m)}}{\bowtie} r_{n,x,m}(\text{student}))$$

(A) names of girl students with the highest marks

(B) names of girl students with more marks than some boy student

(C) names of girl students with marks not less than some boy student

(D) names of girl students with more marks than all the boy students

SOLUTION

This query first computes join where marks are less than or equal to marks are less than or equal to max marks, & then selects the name of all the girl students with more marks than all the boys students. Hence (D) is correct option.

Question. 11

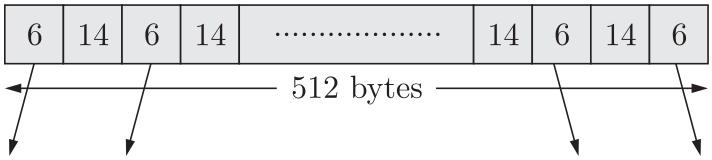
The order of an internal node in a B* tree index is the maximum number of children it can have. Suppose that a child pointer takes 6 bytes, the search field value takes 14 bytes., and the block size is 512 bytes. What is the order of the internal node?

(A) 24 (B) 25

(C) 26 (D) 27

SOLUTION

The structure of a block (node) of B^+ tree



Total if let P pointers than $P-1$ search fields.

$$\begin{aligned}(P-1)14+P\times 6 &= 512\\ 20P-14 &= 512\\ 20P &= 526\\ P &\ \frac{526}{20}\\ &\cong 26\end{aligned}$$

Hence (C) is correct option.

Question. 12

The employee information in a company is stored in the relation

Employee (name, sex, salary, deptName)

Consider the following SQL query

```
select deptname
        from Employee
        where sex='M'
group by deptName
having avg (salary)>
        (select avg(salary)from Employee)
```

It returns the names of the department in which

(A) the average salary is more than the average salary in the company

(B) the average salary of male employees is more than the average salary of all male employees in the company

(C) the average salary of male employees is more than the average salary of employees in the same department

(D) the average salary of made employees is more than the average salary in the company

SOLUTION

"Select average (salary) from employee" selects average salary of the company.

So the query returns the names of all the male employees of all the departments whose average salary is more than the average salary in the company.

Hence (D) is correct option.

YEAR 2005

Question. 13

Which one of the following is a key factor for preferring B^+-trees to binary search trees for indexing database relation?

(A) Database relations have a large number of record

(B) Database relations are sorted on the primary key

(C) B^+-trees require less memory than binary search trees

(D) Data transfer from disks is in blocks

SOLUTION

B^+ tree's each node size is kept almost same to the block size of the system this cause on record data transfer from disk to memory one block at a time.
Hence (D) is correct option.

Question. 14

Which-one of the following statements about normal forms is FALSE?

(A) BCNF is stricter than 3 NF

(B) Loss less, dependency-preserving decomposition into 3 NF is always possible

(C) Loss less, dependency-preserving decomposition into BCNF is always possible

(D) Any relation with two attributes is BCNF

SOLUTION

(A) $BCNF$ is stricter than $3NF$, is true.

(B) True, dependency preserving lossless decomposition into $3NF$ is always possible, by removing transitive functional dependences.

(C) False, lossless dependency preserving decomposition is not possible in all the cases, when all non trivial functional dependencies are not dependent upon super keys.

(D) Certainly relation with two attributes is in $BCNF$, both will consist of super key.

Hence (C) is correct option.

Question. 15

Let r be a relation instance with schema $R = (A, B, C, D)$. WE DEFINE $R_1 = \Pi_{A,B,C}(r)$ and $r_2 = \Pi_{AD}(r)$. let $S = r_1 * r_2$ where $*$ denotes natural join. Given that the decomposition of r into r_1 and r_2 is lossy, which one of the following is TRUE?

(A) $s \subset r$ (B) $r \subset s = r$

(C) $r \subset s$ (D) $r * s = s$

SOLUTION

$r \supset r_1$ & $r \supset r_2$

$s = r_1 * r_2$ * denotes natural join.

So here s would have all the combinations of $r_1 = \Pi_{A,B,C}{}^{(r)}$ & $r_2 = \Pi_{A,D}{}^{(r)}$ product & certainly have more tuples than original r.

So $r \subset s$

Since s would have all the tuples of r but also has some extra.

Hence (C) is correct option.

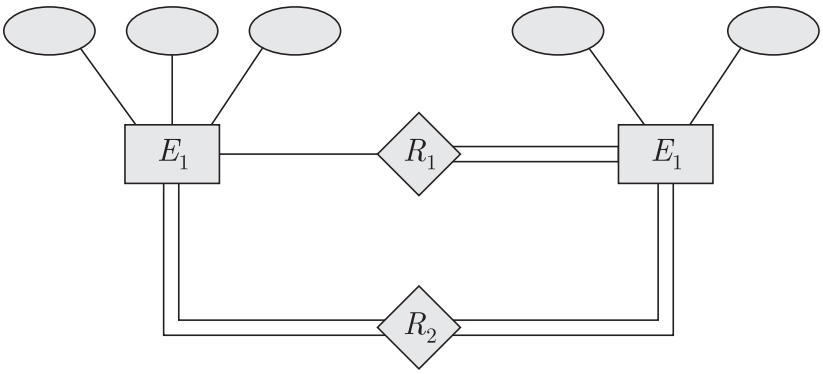
Question. 16

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

(A) 2 (B) 3

(C) 4 (D) 5

SOLUTION



Since R_2 is many to many so E_1 has some primary key so one tables for E_1 & to represent many to many relationship 1 table for R_2 which will have the common attribute of E_2.
R_1 & E_2 can be merged.
So total 3 tables are required.
Hence (B) is correct option.

Question. 17

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

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

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

(A) (3,4) and (6,4) (B) (5,2) and (7,2)

(C) (5,2)(7,2) and (9,5) (D) 1

SOLUTION

On delete cascade says that deletion of a primary key value should delete its all foreign key references.
So in (2,4) 2 is primary key so tuples (5,2) & (7,2) should be deleted, but in (5,2) 5 is also a key so (9,5) also deleted cascade.
Hence (C) is correct option.

Question. 18

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

select title

from book as B

where (select count(*)

from book as T

where T. price>B.Price)<5

(A) Titles of the four most expensive books

(B) Title of the fifth most inexpensive book

(C) Title of the fifth most expensive book

(D) Titles of the five most expensive books

SOLUTION

It selects all those book titles where the price of the book has difference than most expensive book not more than 5.
So it selects five most expensive books.
Hence (D) is correct option.

Question. 19

Consider a relation scheme $R = (A, B, C, D, E, H)$ on which the following functional dependencies hold:

$\{A \rightarrow B,.\ BC \rightarrow D,\ E \rightarrow C, D \rightarrow A\}$

What are the candidate keys of R?

(A) AE, BE (B) AE, BE, DE

(C) AEH, BEH, BCH (D) AEH, BEH, DEH

SOLUTION

$A \rightarrow B,\ BC \rightarrow D,\ E \rightarrow C,\ D \rightarrow A$

We start form set of all the attributes and reduce them using given functional dependences

$ABCDEH$			$ABCDEH$
$ABCEH$	$\{BC \rightarrow D\}$		
$ABEH$	$\{E \rightarrow C\}$	$BCDEH$	$\{D \rightarrow A\}$
AEH	$\{A \rightarrow B\}$	BEH	$\{E \rightarrow C\}$

$ABCDEH$

$ACDEH$ $\{A \rightarrow B\}$

$ADEH$ $\{E \rightarrow C\}$

DEH $\{D \rightarrow A\}$

So candidate keys are AEH, BEH & DEH

Hence (D) is correct option.

YEAR 2006

Question. 20

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

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

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

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

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

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

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

SOLUTION

Here transaction $T1$ has been committed at record 4 so no need to redo them, since balanced commit & start.
This is the step done for recovery.
Hence (C) is correct option.

Question. 21

Consider the relation account (customer, balance) where customer is a primary key and there are no mall values. We would like to

rank customers according to decreasing balance. The customer with the largest balance gets rank 1. Ties are not broken but ranks are skipped: if exactly two customers have the largest balance they each get rank 1 and rank 2 is not assigned.

Query 1 : Select A. customer, count (B. customer) from account A, account B where A. customer

Query 2 : Select A. customer, 1+count(B. customer)from account A, accountB where A, balance<B, balance 7 group by A. customer

Consider these statements about Query 1 and Query 2.

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

Which two of the above statements are correct?

(A) 2 and 5 (B) 1 and 3

(C) 1 and 4 (D) 3 and 5

SOLUTION

Consider an example

Customer	Balance	Correct Rank	Query 1 (Rank)	Query 2 (Rank)
A	40	3	3	3
B	36	4	5	4
C	36	4	5	4
D	44	1	2	1
E	44	1	2	1
			incorrect	correct

Query 1 is incorrect implementation

Query 2 & 1 will give similar result when all balances are distinct, so

statement 1 is correct.
Statement 4 is also correct.
So option (C) is also correct.
More effective query.
Select $A \cdot customer$, $1 + count(B \cdot customer)$ from $amount\, A$, $account\, B$ where $A\, balance < B \cdot balance$ group by $A \cdot balance$

Question. 22

Consider the relation enrolled (student, course) in which student, course) is the primary key, and the relation paid (student, amount) where student is the primary key . Assume no null values and no foreign keys or integrity constraints. Given the following four queries:

Query 1: Select from enrolled where student in (select student form paid)

Query 2: Select student from paid where student in (select student from enrolled)

Query 3: Select E. student from enrolled E, paid P where E. student= P student

Query 4: Se3lect student from paid where exists (select*from enrolled where enrolled student=paid.student

Which one of the following statements is correct?

(A) All queries return identical row sets for any database

(B) Query 2 and Query 4 return identical row sets for all databases but there exist database for which Query 1 and Query 2 retrun different row sets

(C) There exist databases for which Query 3 returns strictly fewer rows than Query 2

(D) There exist databases for which Query 4 will encounter an intergrity violation at runtime

SOLUTION

Query 1 Return student names for different courses, so one student name may occur more than one time.

Query 2 Student listed in paid, one student for one time only.

Query 3 Repeating students.

Query 4 One student selected only once from paid.

So (A) is false.

(B) is correct since (2) & (4) return same rows, but since three is

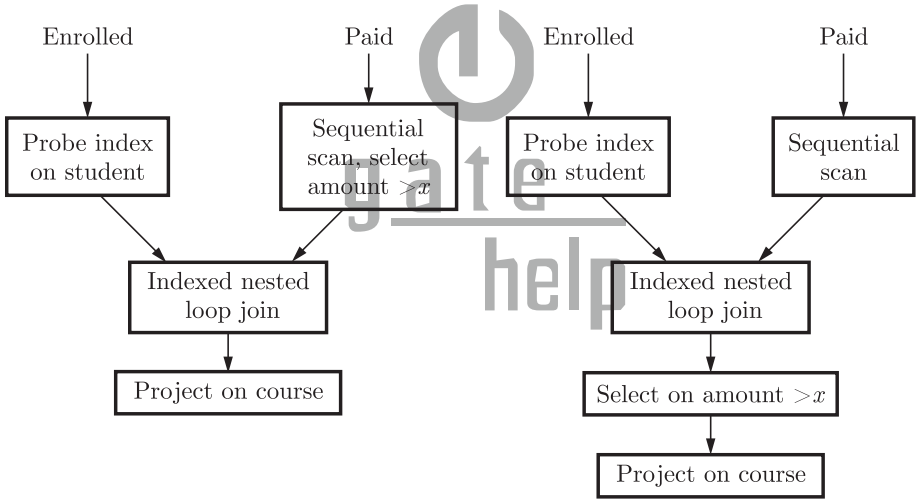
repeating, so even when one student enrolled in one course only then Query 1 and Query 2 results in some rows otherwise their result is always same.

Query 3 can't return few rows than Query 2 in any case.

Hence (B) is correct option.

Question. 23

Consider the relation enrolled (student, course) in which (student, course) is the primary key, and the relation paid (student, amount) where student is the primary key. Assume no null values and no foreign keys or integrity constraints. Assume that amounts 6000, 7000,8000,9000 and 10000 were each paid by 20% of the students. Consider these query plans (Plan 1 on left, Plan 2 on right) to "list all courses taken by students who have paid more than x."



A disk seek takes 4 ms. disk data transfer bank width is 300 MB/s and checking a tuple to see if amount is greater x takes $10\mu s$. Which

of the following statements is correct?

(A) Plan 1 and Plan 2 will not output identical row sets for all databases

(B) A course may be listed more than once int the output of Plan 1 for some databases

(C) For $x = 5000$, Plan 1 executes faster than Plan 2 for all databases

(D) For $x = 9000$, Plan 1 executes slower than Plan 2 for all databases

SOLUTION

Plan 1 selects some tuples which are greater than x, this is best strategy for query evaluation to perform selection as early as possible, to reduce size of the relation.
So plane 1 will execute faster than plan 2 for $x = 5000$.
Since all the tuples required gets selected.
Hence (C) is correct option.

Question. 24

The following functional dependencies are given:

$AB \rightarrow CF, AF \rightarrow D, DE \rightarrow F, C \rightarrow G, F \rightarrow E, G \rightarrow A.$

Which one of the following options is false?

(A) $\{CF\}^+ = \{ACFEFG\}$ (B) $\{BG\}^+ = \{ABCDG\}$
(C) $\{AF\}^+ = \{ACDEFG\}$ (D) $\{AB\}^+ = \{ACDFG\}$

SOLUTION

$AB \rightarrow CD\ AF \rightarrow D\ DE \rightarrow F\ C \rightarrow G\ F \rightarrow E\ G \rightarrow A$
Let us take $\{AF\}^+$
AF
AFE $\{F \rightarrow E\}$
$AFDE$ $\{AF \rightarrow D\}$
$\neq ACDEFG$
Hence (C) is correct option.

YEAR 2007

Question. 25

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

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

(A) Courses in which all the female students are enrolled

(B) Courses in which a proper subset of female students are enrolled

(C) Courses in which only male students are enrolled

(D) None of the above.

SOLUTION

$\Pi_{\text{studId}}(\sigma_{\text{sex}=\text{female}}(\text{studInfo})$ selects student id of all females.
This is joined with all courses (unique) from enroll table.
This results into all the combinations female students can take all the courses.
Difference from enrol cause all the female student id's i.e the sub result eliminated, so virtually whole result is empty.
Hence (A) is correct option.

Question. 26

Consider the relation employee (name, sex, supervisorName (with name as the key. supervisor Name-gives the name of the supervisor of the employee under consideration. What does the following Tuple Relational Calculus query produce?

$\{e.name|employee(e) \vee$

$(\forall \times)[\neg employee(x) \vee x.supervisorName$

$\neq e.name \vee x.sex = \text{"male"}]\}$

(A) Names of employees with a male supervisor

(B) Names of employees with no immediate male subordinates

(C) Names of employees with no immediate female subordinates

(D) Names of employees with a female supervisor

SOLUTION

The sub-query finds employees who are male & have supervisors.
So sub-query results all male employers with at least 1 supervisor and whole query results all those employee names with no immediate female subordinates.
Hence (C) is correct option.

Question. 27

Consider the table employee (empId, name, department, salary) and the two queries Q_1, Q_2 below. Assuming that department 5 has more than one employee, and we want to find the employees who get higher

salary than anyone in the department 5, which one of the statements is TRUE for any arbitrary employee table?

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

$Q._2$: Select e. empId
From employee e
Where e.salary>Any
(Select distinct salary From employee s Where s. department=”5”)

(A) Q_1 is the correct query.
(B) Q_2 is the correct query
(C) Both Q_1 and Q_2 produce the same answer
(D) Neither Q_1 nor Q_2 is the correct query

SOLUTION

Query Q_2 is correct to calculate all the employee *ID's* where salary is greater than any employee of dept no. 5. So correct.
But *Q's* sub-query is opposite also not correct no comparison after where clause.
Hence (B) is correct option.

Question. 28

Which one of the following statements is FALSE?
(A) Any relation with two attributes is in BCNF
(B) A relation in which every key has only one attribute is in 2NF
(C) A prime attribute can be transitively dependent on a key in 3NF relation
(D) A prime attribute can be transitively dependent on a key in a BNCF relation.

SOLUTION

(A) True, since any relation with 2 attributes is in $BCNF$.

(B) True since in $2NF$ non prime attributes should fully functional dependent upon keys.

(C) A prime attribute can functionally dependent on a key in $3NF$ relation, this statement is tree.

(D) But in $BCNF$ transitivity is eliminated.

So false.

Hence (D) is correct option.

Question. 29

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

(A) 63 (B) 64

(C) 67 (D) 68

SOLUTION

Construction of B^+ tree node here.

6 bytes block pointer	7 byte data pointer	9 byte value	7.....	7	9	7

So let P be no. of child pointer then data values $= P-1$

$$\begin{aligned} 6+(P-1)\times 9+P\times 7 &= 1024 \\ 16P-3 &= 1024 \\ 16P &= 1027 \\ P &= \frac{1027}{16} \\ &\cong 64 \end{aligned}$$

Hence (B) is correct option.

Question. 30

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

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

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

(A) Both S_1 and S_2 are conflict serializable

(B) S_1 is conflict serializable and S_2 is not conflict serializable

(C) S_1 is not conflict serializable and S_2 is conflict serializable

(D) Both S_1 and S_2 are not conflict serializable

SOLUTION

Consider S_1 schedule

S_1

$r_1(X) \rightarrow$ read X process 1
$r_1(Y) \rightarrow$ read Y process 1
$r_2(X) \rightarrow$ read X process 2
$r_2(Y) \rightarrow$ read Y process 2
$w_2(Y) \rightarrow$ but this write cause $r_1(Y)$
$w_1(X)$ Value obsolete, so not conflict serializable.

S_2

$r_1(X)$ X read by 1
$r_2(X)$ X read by 2
$r_2(Y)$ Y read by 3
$w_2(Y)$ Here also no problem since only
$r_1(X)$ Transaction 2 has read Y &
$w_1(X)$ It is manipulating
$r_1(Y)$ is after that so conflict serializable.

Hence (C) is correct option.

YEAR 2008

Question. 31

A clustering index is defined on the fields which are of type

(A) Non-key and ordering (B) Non-key and non-ordering

(C) key and ordering (D) Key and non-ordering

SOLUTION

Clustering index is used when we have clusters of similar database enteries for a particular attribute.

So indexing is done by ordering all the instances of attribute & clustering the same values, so index entry for only the first item is needed only.
This attribute has repeating values.
So can't be primary key, so non key.
Hence (A) is correct option.

Question. 32

Let R and S be two relations with the following schema

$R(\underline{P}, \underline{Q}, R1, R2, R3)$

$S(\underline{P}, \underline{Q}, S1, S2)$

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

I $\Pi_P(\bowtie S)$

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

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

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

(A) Only I and II (B) Only I and II

(C) Only I, II and III (D) Only I, II and IV

SOLUTION

I Projects P after joining R & S for common attributes P & Q.

II Selects P from R & P from S & then join so may have same tuples as I.

III Selects P from common PQ attributes & common tuples in R & S

So same as equi join between R & S.

IV Selects P from R which are only in S.

So I, II & III are equivalent.
Hence (C) is correct option.

Question. 33

Consider the following relational schemes for a library database:

Book (Title, Author, Catalog_no, Publisher, Year, price)

Collection (Title, Author, Catalog_no)

Which the following functional dependencies:

I. Title Author →Catalog_no

II. Catalog_no →Title Author Publisher Year

III. Publisher Title Year→price

Assume {Author, Title} is the key for both schemes: which of the following statements is true?

(A) Both Book and Collection are in BCNF

(B) Both Book and Collection are in 3NF only

(C) Book is in 2NF and Collection is in 3NF

(D) Both Book and Collection are in 2NF only

SOLUTION

In relation collection
Title Author → Catalog_no.
So non key attribute Catalog_no. is fully dependent upon key attributes.
So it is in $3NF$.
Relation Book is in $2NF$
Publisher and year attributes are transitively dependent upon title Author, so $3NF$ is not there.
Hence (C) is correct option.

Question. 34

Consider a file of 1684 records. Each record is 32 bytes long and its key field is of size 6 bytes. The file is ordered on a non-key field, and the file organization is unspanned. The file is stored in a file system with block size 1024 bytes, and the size of a block pointer is 10 bytes. If the secondary index is built on the key field of the file, and a multi-level index scheme is used to store the secondary index, the number of first-level and second-level blocks in the multi-level index are respectively

(A) 8 and 0 (B) 128 and 6

(C) 256 and 4 (D) 512 and 5

SOLUTION

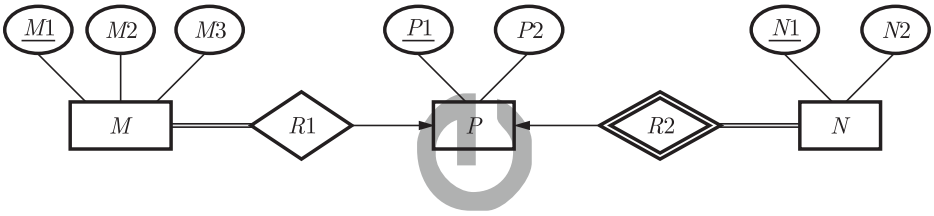
The size of a block = 1024 bytes = 2^{10} bytes
So total of bits required for first level & second level should be 10.
So option (C) $256 = 2^8$
& $4 = 2^2$

Hence (C) is correct option.

Data for Q. 49 & 50 are given below.

Solve the problems and choose the correct answers.

Consider the following ER diagram



Question. 35

The minimum number of tables needed to represent $M, N, P, R1, R2$ is
(A) 2 (B) 3
(C) 4 (D) 5

SOLUTION

There are 3 entity set, but 2 relations R_1 & R_2 only. R_1 is one to many from P to M.
& R_2 is one to many from P to N
N is weak entity set here.
Hence (B) is correct option.

Question. 36

Which of the following is a correct attribute set for one of the tables for the correct answer to the above question?
(A) {M1,M2,M3,P1} (B) {M1,P1,N1,N2}
(C) {M1,P1,N1} (D) {M1,P1}

SOLUTION

Two tables in the scenario will be M_1, M_2, M_3, P_1 which is for relation R_1 & N_1, N_2, P_1, P_2 with P_1 as primary key N_1 is a weak entity set.

This is for relation R_2.

Hence (A) is correct option.

YEAR 2009

Question. 37

Consider two transactions T_1 and T_2 and four schedules S_1, S_2, S_3, S_4 of T_1 and T_2 as given below :

$T1$: $R1\,[x]\ W1\,[x]\ W1\,[y]$

$T2$: $R2\,[x]\ R2\,[y]\ W2\,[y]$

$S1$: $R1\,[x]\ R2\,[x]\ R2\,[y]\ W1\,[x]\ W1\,[y]\ W2\,[y]$

$S2$: $R1\,[x]\ R2\,[x]\ R2\,[y]\ W1\,[x]\ W2\,[y]\ W1\,[y]$

$S3$: $R1\,[x]\ W1\,[x]\ R2\,[x]\ W1\,[y]\ R2\,[y]\ W2\,[y]$

$S4$: $R2\,[x]\ R2\,[y]\ R1\,[x]\ W1\,[x]\ W1\,[y]\ W2\,[y]$

Which of the above schedules are conflict-serializable?

(A) S1 and S2 (B) S2 and S3

(C) S3 only (D) S4 only

SOLUTION

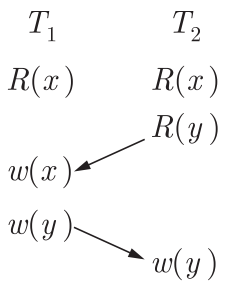
Conflict serializable schedules

Schedule S_1 Schedule S_2

T_1 T_2 T_1 T_2

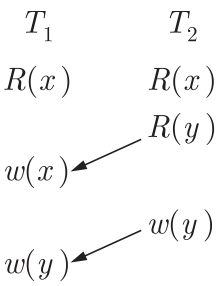
$R(x)$ $R(x)$

Schedule S_1



Cycle exists between T_1 and T_2

Schedule S_2



No Cycle

Cycle exists between T_1 & T_2

No cycle

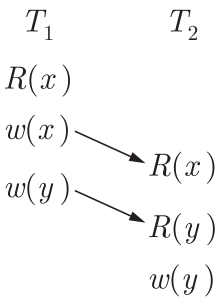
Schedule S_3

T_1 T_2

Schedule S_4

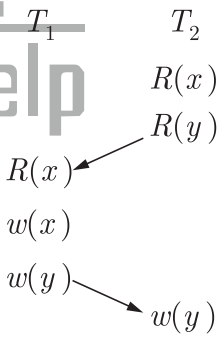
T_1 T_2

Schedule S_3



No cycle

Schedule S_4



Cycle exists

No cycle Cycle exists

In S_2 & S_3 no cycle exists, so conflict serializable.

Hence (B) is correct option.

Question. 38

The following key values are inserted into a $B+$ −tree in which order of the internal nodes is 3, and that of the leaf nodes is 2, in the sequence given below. The order of internal nodes is the maximum number of tree pointers in each node, and the order of leaf nodes is

the maximum number of data items that can be stored in it. The $B+-$tree is initially empty.

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

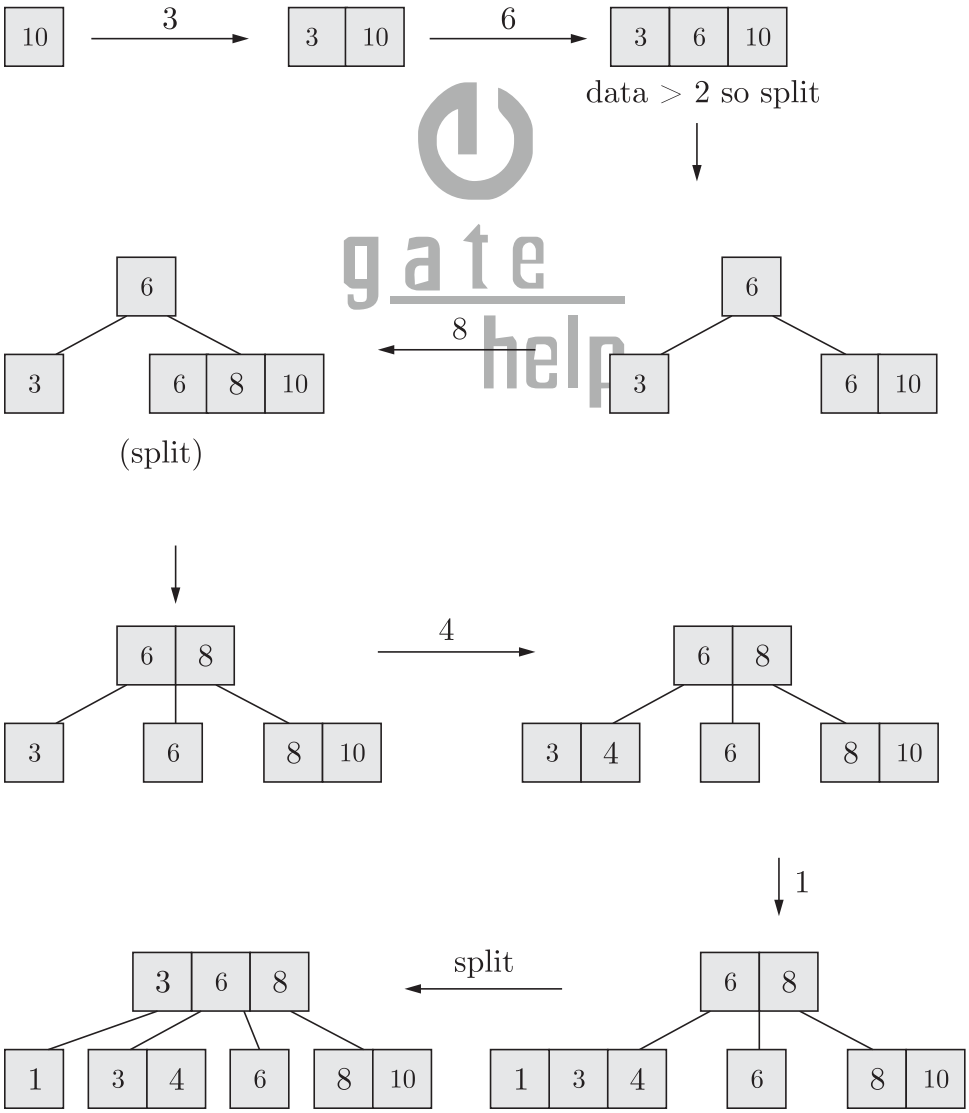
The maximum number of times leaf nodes would get split up as a result of these insertions is

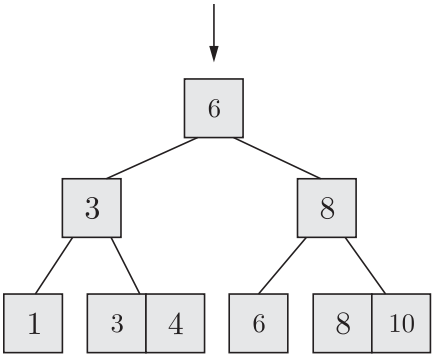
(A) 2 (B) 3

(C) 4 (D) 5

SOLUTION

Order = 3 so any node can have 3 child nodes i.e 3 pointer values.
10, 3, 6, 8, 4, 2, 1





Total 4 splits
Hence (C) is correct option.

Question. 39

Let R and S be relation schemes such that $R = \{a, b, c\}$ and $S = \{c\}$. Now consider the following queries on the database :

I. $\pi_{R-S}(r) - \pi_{R-S}(\pi_{R-S}(r) \times S - \pi_{R-S,S}(r))$

II. $\{t \mid t \in \pi_{R-S}(r) \wedge \forall u \in s(\exists v \in r(u = v[s] \wedge t = v[R - S]))\}$

III. $\{t \mid t \in \pi_{R-S}(r) \wedge \forall v \in r(\exists u \in s(u = v[s] \wedge t = v[R - S]))\}$

IV. Select $R.a, R.b$

From R, S

Where $R.c = S.c$

Which of the above queries are equivalent?

(A) I and II (B) I and III

(C) II and IV (C) III and IV

SOLUTION

Query IV is a natural join between R & S.
The similar result we obtain from Query II.
Hence (C) is correct option.

Common Data for Q. 54 & 55:

Consider the following relational schema:

Suppliers(sid: integer, sname:string, city:string, street:string)

Parts(pid:integer, pname:string, color:string)

Catalog(sid:integer, pid:integer, cost:real)

Question. 40

Consider the following relational query on the above database :

```
SELECT S.sname
FROM Suppliers S
WHERE S.sid NOT IN ( SELECT C.sid
                     FROM Catalog C
                     WHERE  C.pid NOT IN ( SELECT P.pid
                     FROM Parts P
                     WHERE  P.color◇'blue'))
```

Assume that relations corresponding to the above schema are not empty. Which one of the following is the correct interpretation of the above query?

(A) Find the names of all suppliers who have supplied a non-blue part.

(B) Find the names of all suppliers who have not supplied a non-blue part.

(C) Find the names of all suppliers who have supplied only blue parts.

(D) Find the names of all suppliers who have not supplied only blue parts.

SOLUTION

The sub-query results all those parts which are not blue.
The sub-query returns all those suppliers which have the blue parts.
And final return from the query are the supplier id's of all those suppliers which supplied a non blue part.

Question. 41

Assume that, in the suppliers relation above, each supplier and each street within a city has a unique name, and (same, city) forms a candidate key. No other functional dependencies are implied other than those implied by primary and candidate keys. Which one of the following is TRUE about the above schema ?

(A) The schema is in BCNF.

(B) The schema is in 3NF but not in BCNF.

(C) The schema is in 2NF but not in 3NF.

(D) The schema is not in 2NF.

SOLUTION

Sname city forms the candidate key.
Suppliers (sid, sname, city, street)
Sname is primary key.
Sid, sname is also key.
City→street exists.
Street is functionally dependent upon city.
Sname→ city.
So transitivity occurs.
So data base is in $3NF$ but not in $BSNF$.

YEAR 2010

Question. 42

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

(A) 1 (B) 2

(C) 3 (D) 4

SOLUTION

Given the no. of keys $= 5$
So order of any non root node $= 6$
This is maximum order

$$\text{Minimum no. of key} = \left\lceil \frac{P}{2} \right\rceil - 1$$

$$= \left\lceil \frac{5}{2} \right\rceil - 1 = 3 - 1 = 2$$

Hence (B) is correct option

Question. 43

A relational schema for a train reservation database is given below.

Passenger (pid, pname, age)

Reservation (pid, class, tid)

Table : passenger

pid	pname	Age
0	'Sachin'	65
1	'Rahul'	66
2	Sourav'	67
3	'Anil'	69

Table : Reservation

pid	class	tid
0	'AC'	8200
1	'AC'	8201
2	'SC'	8201
3	'AC'	8203
4	'SC'	8204
5	'AC'	8202

What pids are returned by the following SQL query for the above instance of the tables ?

```
SELECT pid
FROM Reservation
WHERE class= 'AC' AND
EXISTS (SELECT*
FROM Passenger
WHERE age>65 AND
Passenger.pid = Reservation.pid)
```

(A) 1,0 (B) 1, 2

(C) 1, 3 (D) 1, 5

SOLUTION

The sub-query places a equity-join between 2 tables and select age >65.
So pid 1, 2 & 3 are selected.
In the main query from 1, 2 & 3 only 1 & 3 have class AC.
Hence (C) is correct option.

Question. 44

Which of the following concurrency control protocols ensure both conflict serializability and freedom from deadlock?

I. 2-phase locking

II Time-stamp ordering

(A) I only (B) II only

(C) Both I and II (D) Neither I nor II

SOLUTION

Two phase locking protocol works by running a transaction in two phases, in first phase locks on various data items is achieved only, once second phase i.e unlocking phase start, no locks can be granted then. So due to waiting to have lock, deadlock may occur.
But time stamp protocol due to consideration of time stamp provide both conflict serializability as well as deadlock avoidance.
Hence (B) is correct option.

Question. 45

Consider the following schedule for transactions T1, T2, and T3 :

```
T1          T2          T3
Read (X)
            Read (Y)
                        Read (Y)
Write (X)
                        Write (X)
            Read (X)
            Write (X)
```

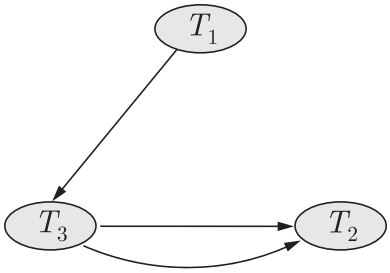
Which one of the schedules below is the correct serialization of the above ?

(A) $T1 \rightarrow T3 \rightarrow T2$ (B) $T2 \rightarrow T1 \rightarrow T3$

(C) $T2 \rightarrow T3 \rightarrow T1$ (D) $T3 \rightarrow T1 \rightarrow T2$

SOLUTION

Drawing the dependency graph where an edge between two transactions exits if there is read write or write-write conflict between them.



Applying topological sort gives order

$$T1 \rightarrow T_3 \rightarrow T2$$

Hence (A) is correct option.

Question. 46

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

$B \rightarrow A$

$A \rightarrow C$

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

(A) 100 (B) 200

(C) 300 (D) 2000

SOLUTION

R has 200 tuples & has 100 tuples, the common attribute for natural join is B in both relations consider a condition cohen all tuples in S has distinct value of B, So a natural join can have maximum 100 tuples.

Hence (A) is correct option.

****

GATE Multiple Choice Questions
For Computer Science

By NODIA and Company

Available in Two Volumes

FEATURES

- The book is categorized into units and the units are sub-divided into chapters.
- Chapter organization for each unit is very constructive and covers the complete syllabus
- Each chapter contains an average of 40 questions
- The questions are standardized to the level of GATE examination
- Solutions are well-explained, tricky and consume less time. Solutions are presented in such a way that it enhances you fundamentals and problem solving skills
- There are a variety of problems on each topic
- Engineering Mathematics is also included in the book