GATE PYQ SERIES
COMPUTER SCIENCE

Databases

Vijay Agarwal sir

Topics To Be Covered

- ER - Model
- Database Design (FD & NF)
- Structured Query Language
- Relational Algebra and Tuple Calculus
- Transactions and Concurrency Control
- File Structures

ER – Model

Q.

Given the basic ER and relational models, which of the following is INCORRECT?

[MCQ: 2012: 1M]

A. An attribute of an entity can have more than one value.

B. An attribute of an entity can be composite.

Multivalued
Composite

C. In a row of a relational table, an attribute can have more than one value.

D. In a row of a relational table, an attribute can have exactly one value or a NULL value.

Ans [C]

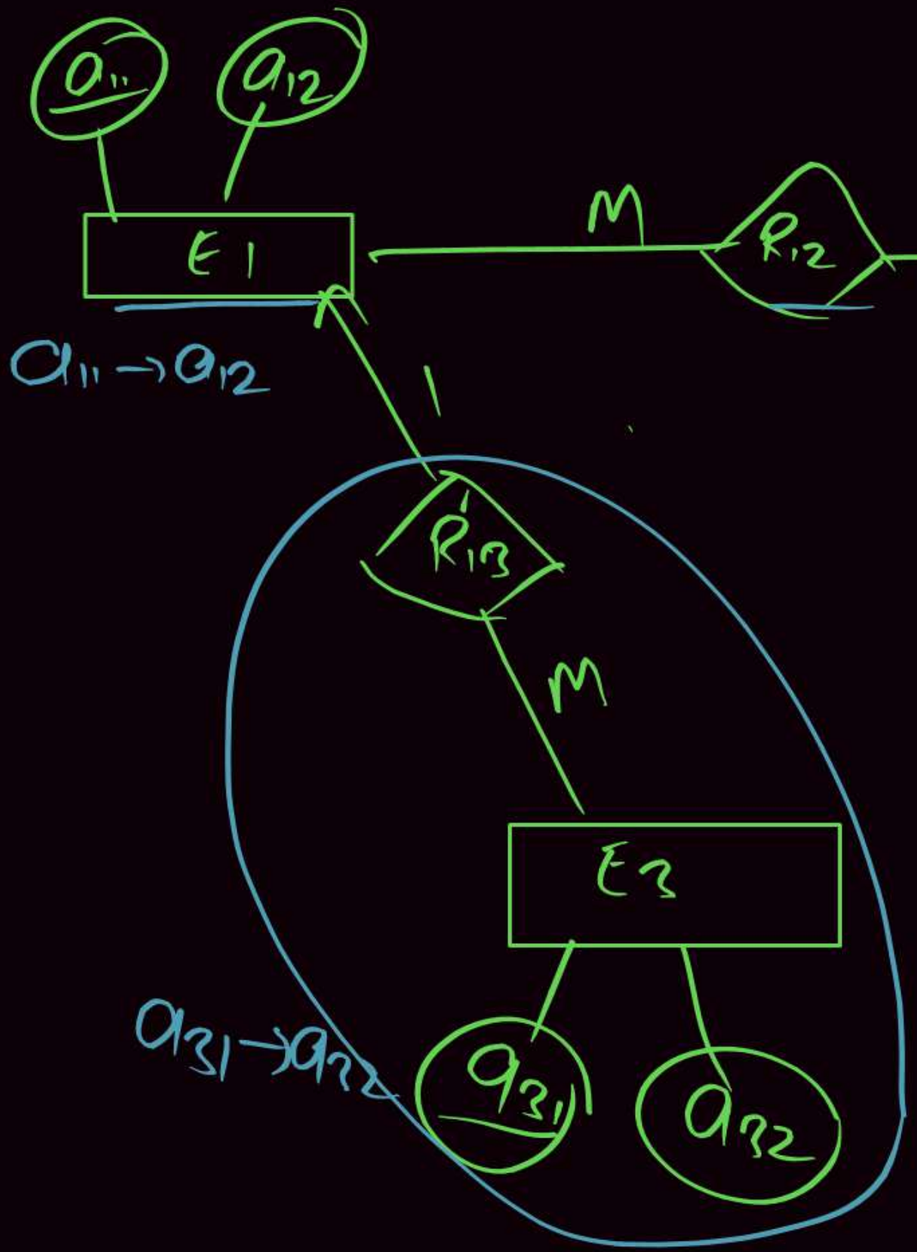
Q.

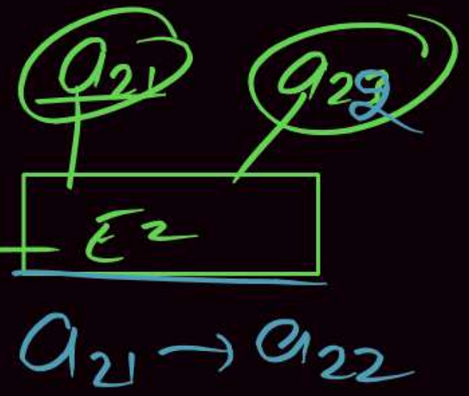
Consider an Entity-Relationship (ER) model in which entity sets E_1 and E_2 are connected by an m:n relationship R_{12}. E_1 and E_3 are connected by a 1 : n (1 on the side of E_1 and n on the side of E_3) relationship R_{13}.

E_1 has two single-valued attributes a_{11} and a_{12} of which a_{11} is the key attribute. E_2 has two single valued attributes a_{21} and a_{22} of which a_{21} is the key attribute. E_3 has two single-valued attributes a_{31} and a_{32} of which a_{31} is the key attribute. The relationships do not have any attributes. If a relational model is derived from the above ER model, then the minimum number of relations that would be generated if all the relations are in 3NF is _____.

4

[NAT: 2015: 2M]





$E_1(\underline{a_{11}}\ a_{12})$ $a_{11} \rightarrow a_{12}$

$E_2(\underline{a_{21}}\ a_{22})$ $a_{21} \rightarrow a_{22}$

$R_{12}(\underline{a_{11}\ b_{11}})$ $\underline{a_{11}\ b_{11}}$ F.K

$E_3R_{13}(\underline{a_{31}\ a_{32}\ a_{11}})$

$a_{31} \rightarrow a_{32}a_{11}$

Q.

An ER model of a database consists of entity types A and B. These are connected by a relationship R which does not have its own attribute. Under which one of the following conditions, can the relational table for R be merged with that of A?

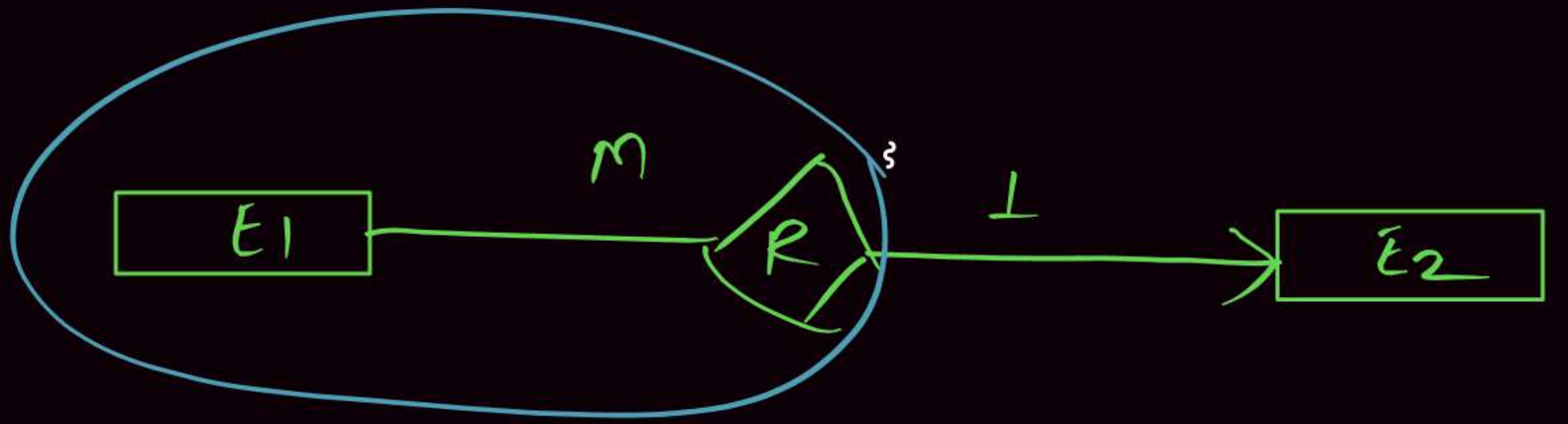
[MCQ: 2017: 1M]

A. Relationship R is one-to-many and the participation of A in R is total.

B. Relationship R is one-to-many and the participation of A in R is partial.

C. Relationship R is many-to-one and the participation of A in R is total.

D. Relationship R is many-to-one and the participation of A in R is partial.


M
S
E1
R
1
E2

Q.

In an Entity-Relationship (ER) model, suppose R is a many-to-one relationship from entity set E1 to entity set E2. Assume that E1 and E2 participate totally in R and that the cardinality of E1 is greater than the cardinality of E2. Which one of the following is true about R?

[MCQ: 2018: 1M]

A. Every entity in E1 is associated with exactly one entity in E2.

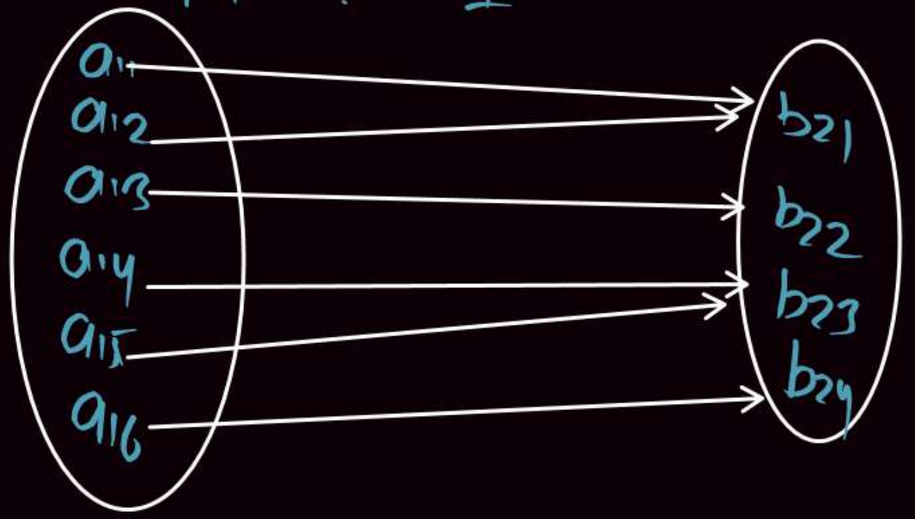
B. Some entity in E1 is associated with more than one entity in E2.

C. Every entity in E2 is associated with exactly one entity in E1.

D. Every entity in E2 is associated with at most one entity in E1.

Coordinality of $E_1 > E_2$

M : 1



Q.

Which one of the following is used to represent the supporting many-one relationships of a weak entity set in an entity-relationship diagram?

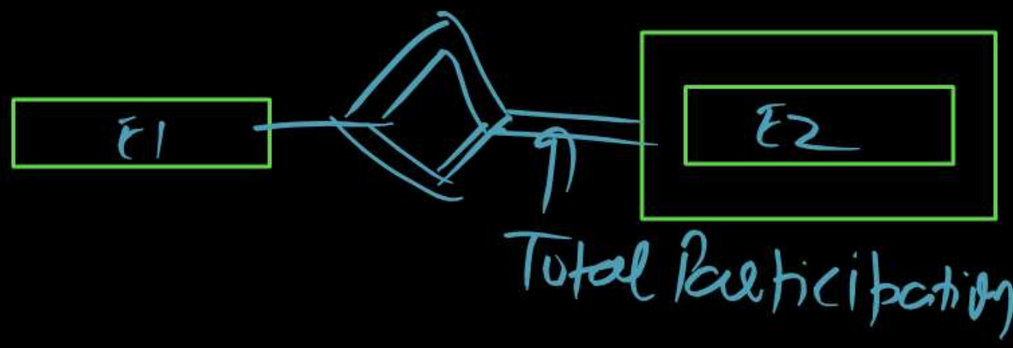
[MCQ: 2020-1M]

A. Ovals that contain underlined identifiers.

B. Diamonds with double/bold border.

C. Ovals with double/bold border.

D. Rectangles with double/bold border.



Database Design

Q.

Common Data for Next two Question

Relation R has eight attributes ABCDEFGH. Fields of R contain only atomic values.

$F = \{CH \rightarrow G, A \rightarrow BC, B \rightarrow CFH, E \rightarrow A, F \rightarrow EG\}$ is a set of functional dependencies (FDs) so that F^+ is exactly the set of FDs that hold for R.

AD
ED
FD
BD
Ans

Prime Attribute = [D, A, E, F, B]

How many candidate keys does the relation R have? [MCQ: 2013: 2M]

A 3

B 4

C 5

D 6

$(AD)^+ = [ABCDEFGH]$

AD is Candidate key

If $X_{Attribute} \rightarrow$ [Prime Attribute]

$E \rightarrow A$

ED is C.K — (2)

$F \rightarrow EG$

FD is C.K — (3)

$B \rightarrow CFH$

BD is C.K — (4)

First find Any One Candidate key

If $X_{Attribute} \longrightarrow$ {Prime Attribute}

then Multiple Candidate keys are there

Q.

Relation R has eight attributes ABCDEFGH. Fields of R contain only atomic values.

$F = \{CH \rightarrow G, A \rightarrow BC, B \rightarrow CFH, E \rightarrow A, F \rightarrow EG\}$ is a set of functional dependencies (FDs) so that F^+ is exactly the set of FDs that hold for R.

[MCQ: 2013: 2M]

The relation R is

(A) in 1NF, but not in 2NF.

(B) in 2NF, but not in 3NF.

(C) in 3NF, but not in BCNF.

(D) in BCNF.

Q.

Given an instance of the STUDENTS relation as shown below:

StudentID	StudentName	StudentEmail	StudentAge	CPI
2345	Shankar	shankar@math	X	9.4
1287	Swati	swati@ee	19	9.5
7853	Shankar	shankar@cse	19	9.4
9876	Swati	swati@mech	18	9.3
8765	Ganesh	ganesh@civil	19	8.7

For (StudentName, StudentAge) to be a key for this instance, the value X should NOT be equal to 19.

[NAT: 2014: 1M]

Q. The maximum number of superkeys for the relation schema R(E, F, G, H) with E as the key is 8. [NAT: 2014: 1M]

R(EFGH)

Candidate key = E

#Superkey = E FGH

$2^3 = 8$ Superkey

(OR)

#Superkey = $2^{4-1} = 2^3 = 8$ Superkey

E FGH

000 → E
001 → EH
010 → EG
011 → EGH
100 → EF
101 → EFH
110 → EFG
111 → EFGH

Q. A prime attribute of a relation scheme R is an attribute that appears

[MCQ: 2014: 1M]

A. in all candidate keys of R.

B. in some candidate key of R.

C. in a foreign key of R.

D. only in the primary key of R.

Prime Attribute : [A, B, C]

R(ABCDE) [A→BCD , D→E, BC→AD]

$(A)^+ = [ABCDE]$

A is Candidate key ①

BC → [A]

$(BC)^+ = [BCADE]$

$(B)^+ = [B]$

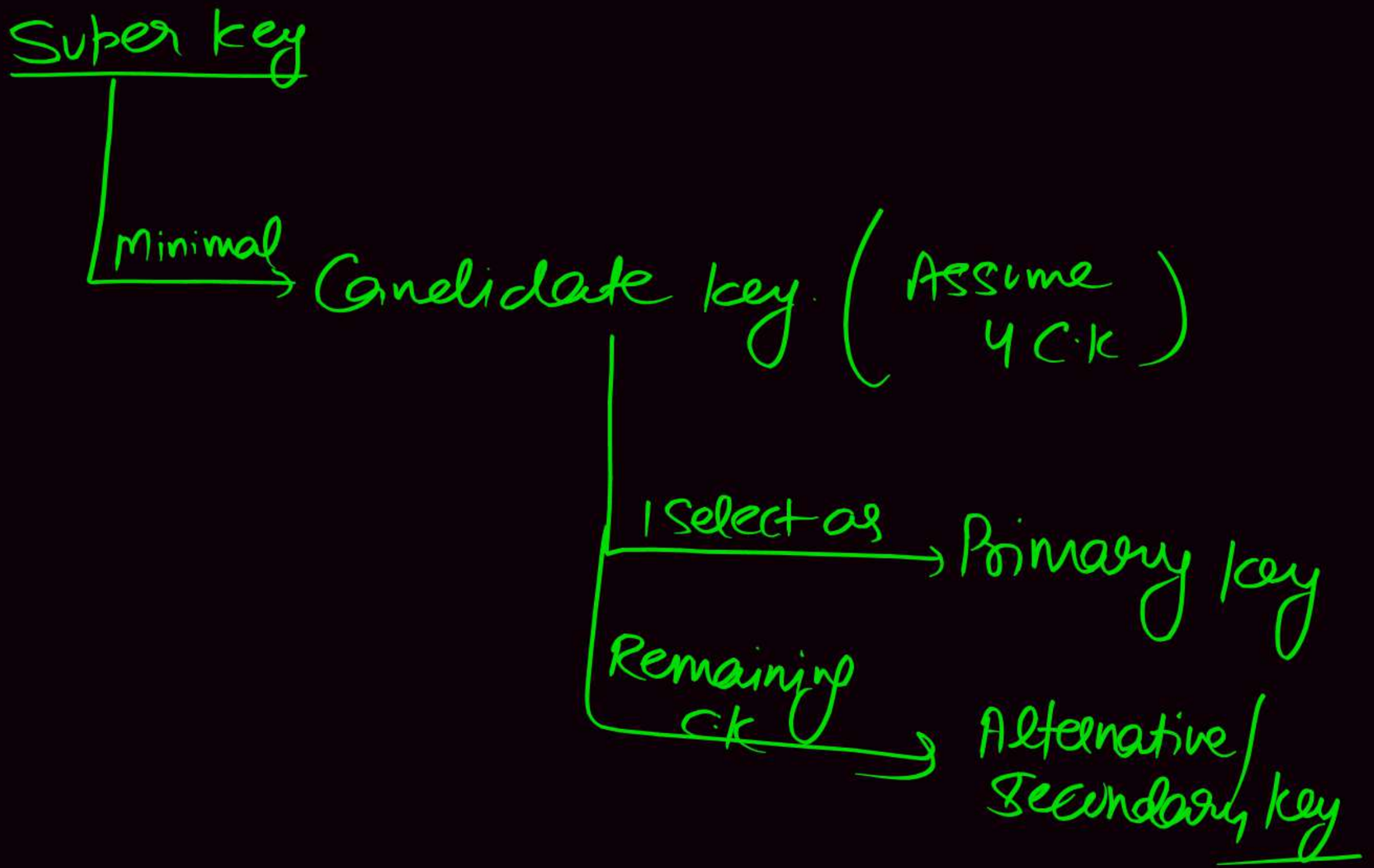
$(C)^+ = [C]$

BC is Candidate key ②

R(ABCD) [A → B, B → C, C → D, D → A]

Candidate key = [A, D, C, B]

Prime Attribute = [A, B, C, D]


Super key
Minimal
Candidate key (Assume 4 C.K)
1 Select as
Primary key
Remaining C.K
Alternative / Secondary key

Q.

Given the following two statements:

S1: Every table with two single-valued attributes is in 1NF, 2NF, 3NF and BCNF.

S2: $AB \rightarrow C, D \rightarrow E, E \rightarrow C$ is a minimal cover for the set of functional dependencies $AB \rightarrow C, D \rightarrow E, AB \rightarrow E, E \rightarrow C$.

Which one of the following is CORRECT?

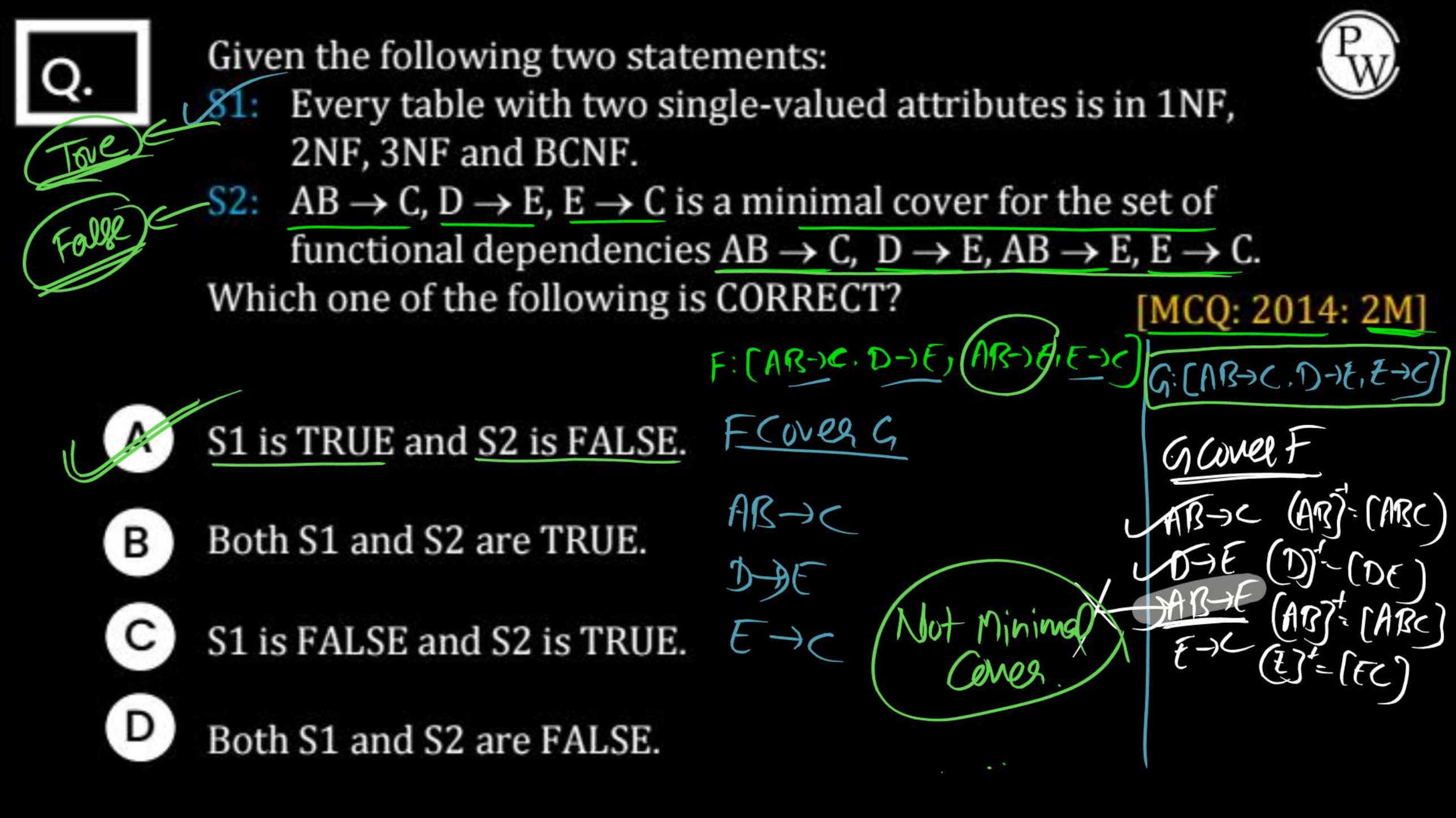
[MCQ: 2014: 2M]

(A) S1 is TRUE and S2 is FALSE.

(B) Both S1 and S2 are TRUE.

(C) S1 is FALSE and S2 is TRUE.

(D) Both S1 and S2 are FALSE.



Minimal Cover

F: [ ]

G: [ ]

R(AB) [A → B]

C.K : A

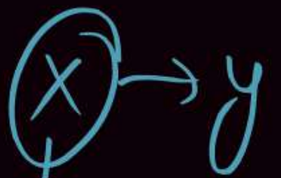
R(AB) [B → A]

C.K : B

R(AB) [A → B, B → A]

C.K : A, B

BCNF



Super key

Q.

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

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

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

I (VOLUME, NUMBER, STARTPAGE, ENDPAGE) → TITLE

II (VOLUME, NUMBER) → YEAR → P.D Not in 2NF.

III (VOLUME, NUMBER, STARTPAGE, ENDPAGE) → PRICE.

The database is redesigned to use the following schemas.

BCNF ← (VOLUME, NUMBER, STARTPAGE, ENDPAGE, TITLE, PRICE) ✓ I & III

BCNF ← (VOLUME, NUMBER, YEAR)

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

[MCQ: 2016: 1M]

Volume Number → year

BCNF.

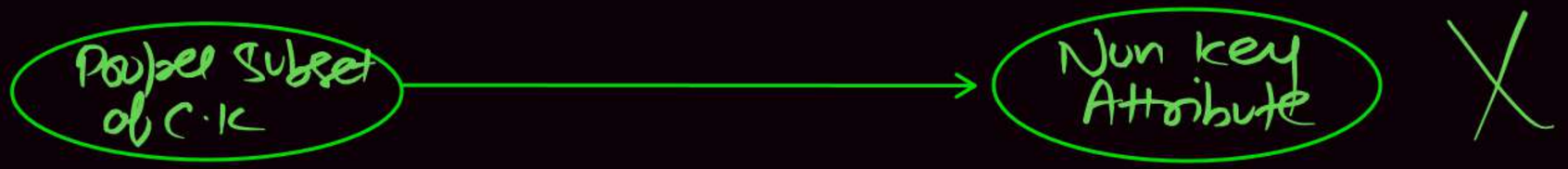
(A) 1NF

(B) 3NF

(C) 2NF

(D) BCNF

2NF Partial Dependency



Q. Which of the following is NOT a superkey in a relational schema with attributes V, W, X, Y, Z and primary key V Y?

[MCQ: 2016: 1M]

A VXYZ

B VWXZ

C VWXY

D VWXYZ

VY

Any Super Set of VY is a Key

Q.

The following functional dependencies hold true for the relational schema R{V, W, X, Y, Z}:

$V \rightarrow W$
$VW \rightarrow X$
$Y \rightarrow VX$
$Y \rightarrow Z$

Which of the following is irreducible equivalent for this set of functional dependencies? [MCQ:2017: 1M]

A
$V \rightarrow W$
$V \rightarrow X$
$Y \rightarrow V$
$Y \rightarrow Z$

B
$V \rightarrow W$
$W \rightarrow X$
$Y \rightarrow V$
$Y \rightarrow Z$

C
$V \rightarrow W$
$V \rightarrow X$
$Y \rightarrow V$
$Y \rightarrow X$
$Y \rightarrow Z$

D
$V \rightarrow W$
$W \rightarrow X$
$Y \rightarrow V$
$Y \rightarrow X$
$Y \rightarrow Z$

$V \to W$ $\quad VW \to X$ $\quad Y \to VX$ $\quad Y \to Z$

Step1

$V \to W$	$VW \to X$	$Y \to V$	$Y \to X$	$Y \to Z$

Bind Redundant FD

Step2

$V \to W$ ✓ $\quad VW \to X$ ✓ $\quad Y \to V$ ✓

$Y \to X$

$(Y)^+ = [Y\,V\,Z\,W\,X]$

$Y \to Z$

$(Y)^+ = [Y\,V\,W\,X]$

$(V)^+ = [V]$ $\quad (VW)^+ = [VW]$ $\quad (Y)^+ = [Y\,X\,Z]$

Step2

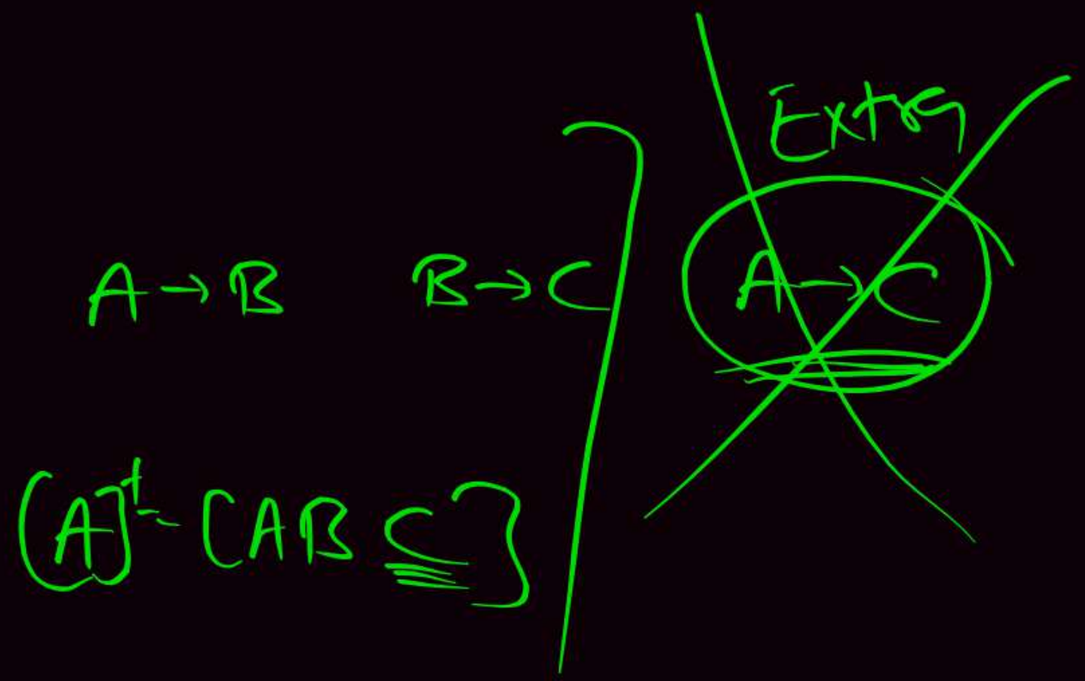
$V \to W$	$VW \to X$	$Y \to V$	$Y \to Z$

$(V)^+ = [VW]$ $\quad (W)^+ = [W]$ $\quad \cancel{VW \to X}$

W is extra if $(V)^+$ = Contain W

V is extra if $(W)^+ = [\ldots \cancel{V}]$

$V \to W$
$V \to X$
$Y \to V$
$Y \to Z$


A → B
B → C
Extra
A → C
(A)+ CAB

Minimal Cover

① Split the FD Such that RHS Contain Single Attribute

② Find the Redundant FD & Delete them from FD set

③ Find Extra Attribute on L.H.S

Q.

Consider a relational table R that is in 3NF, but not in BCNF. Which one of the following statements is TRUE?

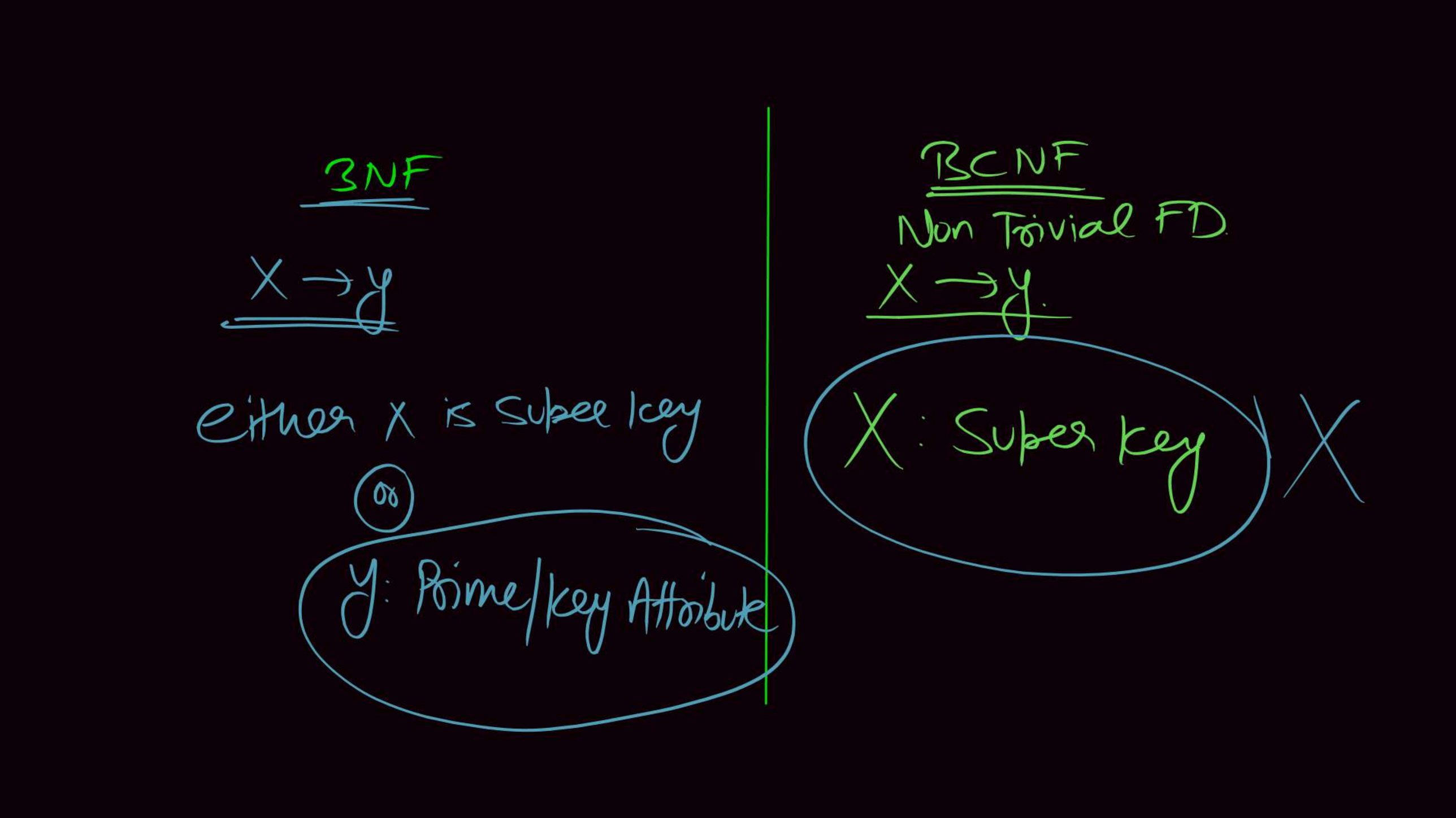
[MCQ: 2020-2M]

A. R has a nontrivial functional dependency X→A, where X is not a superkey and A is a non-prime attribute and X is a proper subset of some key.

B. R has a nontrivial functional dependency X→A, where X is not a superkey and A is a non-prime attribute and X is not a proper subset of any key.

C. A cell in R holds a set instead of an atomic value.

D. R has a nontrivial functional dependency X→A, where X is not a superkey and A is a prime attribute.


3NF
X → Y
either X is Super key
(or)
Y: Prime/key Attribute
BCNF
Non Trivial FD.
X → Y.
X: Super key
X

Q.

Consider the relation R (P, Q, S, T, X, Y, Z, W) with the following functional dependencies.

$$PQ \rightarrow X;\ P \rightarrow YX;\ Q \rightarrow Y;\ Y \rightarrow ZW$$

Consider the decomposition of the relation R into the constituent relations according to the following two decomposition schemes.

D_1: $R = [(P, Q, S, T); (P, T, X); (Q, Y); (Y, Z, W)]$

D_2: $R = [(P, Q, S); (T, X); (Q, Y); (Y, Z, W)]$

Which one of the following options is correct?

[MCQ: 2021: 2M]

(A) D_1 is a lossless decomposition, but D_2 is a lossy decomposition.

(B) D_1 is a lossy decomposition, but D_2 is a lossless decomposition.

(C) Both D_1 and D_2 are lossless decomposition.

(D) Both D_1 and D_2 are lossy decomposition.

D2: $R_1(PQS)$ $R_2(TX)$ $R_3(QY)$ $R_4(YZW)$

FD: $[PQ \to X,\ P \to YX,\ Q \to Y,\ Y \to ZW]$

$R_1(PQS) \cap R_3(QY) \Rightarrow Q \Rightarrow [Q]^+ = [QY\ldots]$ super key of R_3

$R_{13}(PQSY) \cap R_4(YZW) \Rightarrow Y \Rightarrow [Y]^+ = [YZW)$ super key of R_4

$R_{134}(PQSYZW) \cap R_2(TX)$

No Common Attribute

D1: $R_1(PQST)$ $R_2(PTX)$ $R_3(Qy)$ $R_4(yzw)$

FD: $\{PQ \to X,\ P \to yX,\ Q \to y,\ y \to zw\}$

$R_1(PQST) \cap R_2(PTX) \Rightarrow PT \Rightarrow [PT]^+ = [PTyX\ldots]$ Super key of R_2

$R_{12}(PQSTX) \cap R_3(Qy) \Rightarrow [Q] \Rightarrow [Q]^+ = [Qy\ldots]$ Super key of R_3

$R_{123}(PQSTXy) \cap R_4(yzw) \Rightarrow y \Rightarrow (y)^+ = [yzw]$ Super key of R_4

$R_{1234}(PQSTXyzw)$

D1: Lossless

① $R_1 \cup R_2 \equiv R$

Lossless Join

② $R_1 \cap R_2$

If Common Attribute

either super key of R_1

(or)

super key of R_2.

$R_1 \cap R_2 \longrightarrow R_1$

(or)

$R_1 \cap R_2 \longrightarrow R_2$

Lossless

Q.

Consider the following statements S1 and S2 about the relational data model:

S1: A relation scheme can have at most one foreign key.

S2: A foreign key in a relation scheme R cannot be used to refer to tuples of R.

Which one of the following choices is correct?

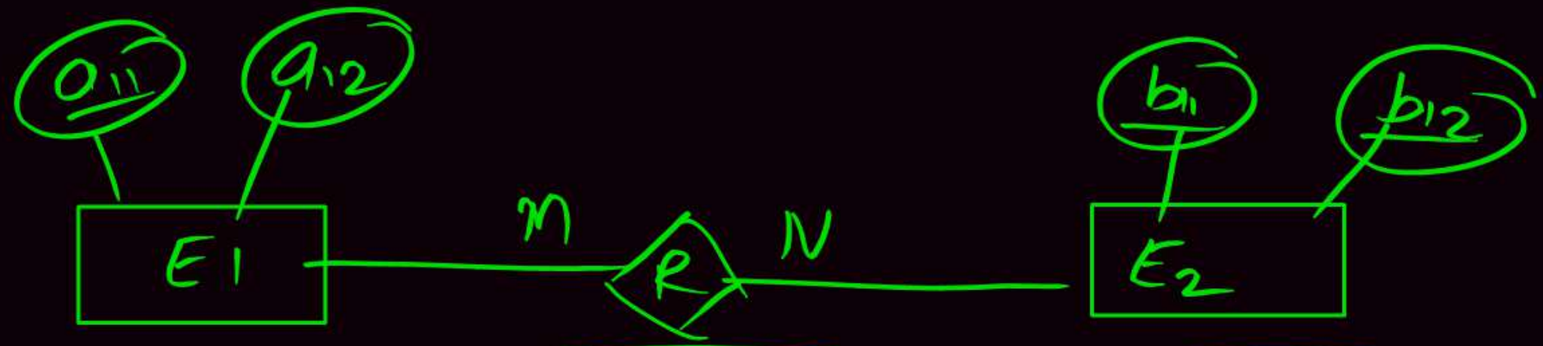
[MCQ: 2021 – 1M]

A. Both S1 and S2 are true.

B. S1 is true and S2 is false.

C. S1 is false and S2 is true.

D. Both S1 and S2 are false.



S_1:

False

$E_1(\underline{a_{11}}\ a_{12})$ $R(\underline{a_{11}\ b_{11}})$ $E_2(\underline{b_{11}}\ b_{12})$

Emp

eid	ename	manager
E1		E_3
E_2		E_3
E_3		E_3
E_4		E_1

Q.

Suppose the following functional dependencies hold on a relation U with attributes P, Q, R, S and T:

$P \rightarrow QR$

$RS \rightarrow T$

Which of the following functional dependencies can be inferred from the above functional dependencies? [MSQ: 2021 - 2M]

A. $PS \rightarrow T$

B. $R \rightarrow T$

C. $P \rightarrow R$

D. $PS \rightarrow Q$

$[PS]^+ = [PSQRT]$

$[R]^+ = [R]$

$[P]^+ = [PQR]$

A, C and D

Q. In a relational data model, which one of the following statements is TRUE?

[MCQ: 2022: 1M]

A. A relation with only two attributes is always in BCNF.

B. If all attributes of a relation are prime attributes, then the relation is in BCNF.

C. Every relation has at least one non-prime attribute.

D. BCNF decompositions preserve functional dependencies.

A

R(ABCD) [AB→CD , D→A)

Candidate key = [AB , DB)

Not in BCNF : D → A

Not Super key

BCNF Decomposition

D → A

R(ABCD)

R_2(DA)

AB → CD

Not Preserved

R_1(BCD)

R_2(DA)

R(ABCDE) [A→B, B→C, C→D, D→A]

Candidate key = [AE, BE, CE, DE]

Prime Attribute = [A, B, C, D, E]

3NF But Not in BCNF.

R(AB) [A→B] | [B→A] | [A→B . B→A]

X: super key.

Q.

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

$AB \rightarrow C$; $BC \rightarrow D$; $C \rightarrow E$;

The number of superkeys in the relation R is 8.

[NAT:2022-1M]

$F: [AB \rightarrow C,\ BC \rightarrow D,\ C \rightarrow E)$

$[AB]^+ = [ABCDE)$

$(A)^+ = [A]$

$(B)^+ = [B]$

AB is Candidate key

#Super key $= 2^{5-2} \Rightarrow 2^3 = 8$

$R(ABCDE)$

$\Downarrow$

$2^3 \Rightarrow$ 8 S.K

Structured Query Language

Q.

Database table by name Loan_Records is given below:
What is the output of the following SQL query?

```
SELECT count(*)
FROM (
        SELECT Borrower, Bank_Manager
        FROM Loan_Records) AS S NATURAL JOIN
        (SELECT Bank-Manager, Loan_Amount
         FROM Loan_Records) AS T
);
```

[MCQ: 2011:2M]

A 3

B 9

C 5

D 6

Borrower	Bank_Manager	Loan_Amaount
Ramesh	Sunderajan	10000.00
Suresh	Ramgopal	5000.00
Mahesh	Sunderajan	7000.00

S

Rname	B.m
Ramesh	SJ
Suresh	RG
mahesh	SJ

3 Tuple

2 Attribute

T

BM	L.A
SJ	10k
RG	5k
SJ	7k

3 Tuple

2 Attribute

3 x 3 = 9 Tuple

2 + 2 = 4 Attribute

S.B	S.BM	T.BM	T.LA
Ramesh	SJ	SJ	10k
Ramesh	SJ	RG	5k
Ramesh	SJ	SJ	7k
Suresh	RG	SJ	
Suresh	RG	RG	
Suresh	RG	SJ	
Mahesh	SJ	SJ	
Mahesh	SJ	RG	
Mahesh	SJ	SJ	

Q.

Consider the following relation

Cinema (theatre, address, capacity)

Which of the following options will be needed at the end of the SQL query

SELECT P1.address

FROM Cinema P1

such that it always finds the addresses of theaters with maximum capacity?

[MCQ: 2015-1M]

A. WHERE P1.capacity > = ALL (select P2.capacity from Cinema P2)

B. WHERE P1.capacity > = Any (select P2.capacity from Cinema P2)

C. WHERE P1.capacity > ALL (select max (P2.capacity) from Cinema P2)

D. WHERE P1.capacity > Any (select max (P2.capacity) from Cinema P2)

400

ALL $\geq$ $(x \geq 100)$ AND $(x \geq 200)$ AND $(x \geq 300)$ AND $(x \geq 400)$

Thread	Address	Capacity
T_1	A	100
T_2	B	200
T_3	C	300
T_4	D	400

Q.

Consider the following relations:

Student (S)	
Roll_No	Student_Name
1	Raj
2	Rohit
3	Raj

Performance (P)		
Roll_No	Course	Marks
1	Math	80
1	English	70
2	Math	75
3	English	80
2	Physics	65
3	Math	80

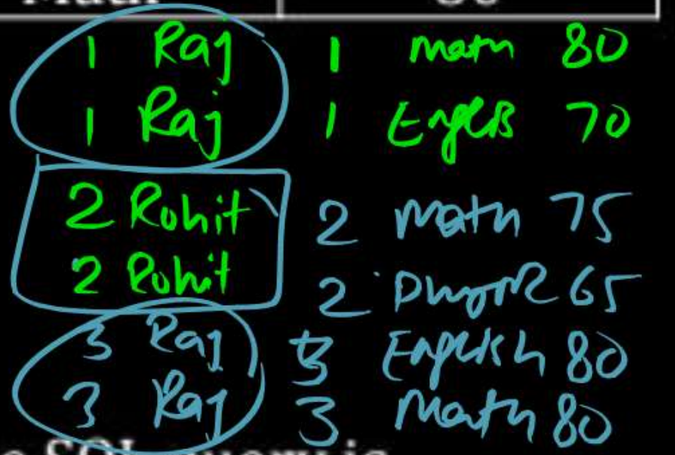
Name	Sum
Raj	310
Rohit	140

Consider the following SQL query.

```
SELECT S . Student_Name, Sum(P. Marks)
FROM Student S, Performance P
WHERE S. Roll_No = P.Roll_No
Group BY S. Student_Name
```

The number of rows that will be returned by the SQL query is _____.

[NAT: 2015–2M]



Q. SELECT operation in SQL is equivalent to [MCQ: 2015-1M]

A The selection operation in relational algebra.

B The selection operation in relational algebra, except that SELECT in SQL retains duplicates.

C The projection operation in relational algebra.

D The projection operation in relational algebra, except that SELECT in SQL retains duplicates.

SELECT (DISTINCT) $A_1 A_2 \ldots A_n$ ≡ Projection(π)

FROM $R_1 R_2 \ldots R_m$ ≡ CROSS Product(X)

WHERE Condition ≡ Selection[σ]

$$\pi_{A_1 A_2 \ldots A_n} \left[\sigma_{\text{Condition}} (R_1 \times R_2 - - - \times R_m) \right]$$

Q.

Consider the following database table named water_schemes:
The number of tuples returned by the following SQL query is:

100/4 = 25

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

Total

Name	Capacity
Ajmer	20
BKNR	40
Churu	30
DNGR	10

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

Total Avg

Avg
25

2

[NAT:2016-2M]

Q.

Consider a database that has the relation schema EMP(EmpId, EmpName, and DeptName). An instance of the schema EMP and a SQL query on it are given below:

```
SELECT AVG(EC.Num)
FROM EC
WHERE(DeptName, Num) IN
        (SELECT DeptName, COUNT(EmpId)AS
                                EC(DeptName, Num)
        FROM EMP
        GROUP BY DeptName)
```

The output of executing the SQL query is _____.

[NAT:2017-1M]

EMP		
EmpId	EmpName	DeptName
1	XYA	AA
2	XYB	AA
3	XYC	AA
4	XYD	AA
5	XYE	AB
6	XYF	AB
7	XYG	AB
8	XYH	AC
9	XYI	AC
10	XYJ	AC
11	XYK	AD
12	XYL	AD
13	XYM	AE

EC

DeptNo	Count(Empid)
AA	4
AB	3
AC	3
AD	2
AE	1

13/5 = 2.6

Q.

Consider the following database table named top_scorer.

Consider the following SQL Query:

```
SELECT ta.player FROM top_scorer AS ta
WHERE ta.goals > ALL (SELECT tb.goals
    FROM top_scorer AS tb
    WHERE tb.country = 'Spain')
AND ta.goals > ANY (SELECT tc.goals
    FROM top_scorer AS tc
    WHERE tc.country = 'Germany')
```

The number of tuples returned by the above SQL query is 7.

[NAT:2017–2M]

top_scorer		
player	country	goals
Klose	Germany	16
Ronaldo	Brazil	15
G Muller	Germany	14
Fountaine	France	13
Pele	Brazil	12
Klinsmann	Germany	11
Kocsis	Hungary	11
Batistuta	Argentina	10
Cubillas	Peru	10
Lato	Poland	10
Lineker	England	10
T Muller	Germany	10
Rahn	Germany	10

ANY ⇒ OR

x > ANY (10, 20, 30)

x > 10 OR x > 20 OR x > 30

O/P → 11, 12, 13, 14, 15, 16

x < ANY (10, 20, 30)

(x < 10) OR (x < 20) OR (x < 30)

O/P → 29, 28, 27, 26, 25

x > ALL (10, 20, 30)

(x > 10) AND (x > 20) AND (x > 30)

O/P → 31, 32, 33 . . .

x < ALL (10, 20, 30)

(x < 10) AND (x < 20) AND (x < 30)

O/P → 9, 8, 7, 6, 5

Q.

Consider the following four relational schemas. For each schema, all non-trivial functional dependencies are listed. The underlined attributes are the respective primary keys.

Schema I: Registration (rollno, courses)

Field 'courses' is a set-valued attribute containing the set of courses a student has registered for.

Non-trivial functional dependency:

rollno → courses ⟶ BCNF

Schema II: Registration (rollno, courseid, email)

Non-trivial functional dependencies:

rollno, courseid → email

email → rollno ⟶ P.A

3NF But Not in BCNF

Schema III: Registration (rollno, courseid, marks, grade)

Non-trivial functional dependencies:

rollno, courseid → marks, grade

marks → grade (N.K ⟶ N.K)

2NF But Not in 3NF.

Schema IV: Registration (rollno, courseid, marks, credit)

Non-trivial functional dependencies:

rollno, courseid → credit

courseid → credit } P.D } Not in 2NF

Which one of the relational schemas above is in 3NF but not in BCNF? [MCQ: 2018: 2M]

(A) Schema I

(B) Schema II

(C) Schema III

(D) Schema IV

Q.

Consider the following two tables and four queries in SQL:
Book(isbn, bname), Stock(isbn, copies)

Query 1:
```
SELECT B.isbn, S.copies
FROM Book B INNER JOIN Stock S ON B.isbn = S.isbn;
```

Query 2:
```
SELECT B.isbn, S.copies
FROM Book B LEFT OUTER JOIN Stock S ON B.isbn = S.isbn;
```

Query 3:
```
SELECT B.isbn, S.copies
FROM Book B RIGHT OUTER JOIN Stock S ON B.isbn = S.isbn;
```

Query 4:
```
SELECT B.isbn, S.copies
FROM Book B FULL OUTER JOIN Stock S ON B.isbn = S.isbn;
```

[MCQ: 2018: 1M]

Which one of the queries above is certain to have an output that is superset of the outputs of the other three queries?

(A) Query 1
(B) Query 3
(C) Query 2
(D) Query 4

Natural Join

Left Outer JOIN ⇒ $\bowtie$ && Left

Right outer Join ⇒ $\bowtie$ && Right

FULL OUTER JOIN ⇒ LOJ U ROJ

Q.

A relational database contains two tables Student and Performance as shown below:

Amit 452
Priya 452
Vinit 452
Rohan 452
Smita 452

Student [S]

Roll_no	Student_name
1	Amit
2	Priya
3	Vinit
4	Rohan
5	Smita

Performance [P]

Roll_no	Student_code	Marks
1	A	86
1	B	95
1	C	90
2	A	89
2	C	92
3	C	80

The primary key of the Student table is Roll_no. For the Performance table, the columns Roll_no. and Subject_code together from the primary key. Consider the SQL query given below:

```
SELECT S.Student_name, sum(P.Marks) FROM Student S, Performance P
WHERE P. Marks > 84 GROUP BY S. Student_name;
```

The number of rows returned by the above SQL query is ____.

5

[NAT: 2019–2M]

Q.

Consider a relational database containing the following schemas.

Catalogue

sno	pno	cost
S1	P1	150
S1	P2	50
S1	P3	100
S2	P4	200
S2	P5	250
S3	P1	250
S3	P2	150
S3	P5	300
S3	P4	250

Suppliers

sno	sname	location
S1	M/s Royal furniture	Delhi
S2	M/s Balaji furniture	Bangalore
S3	M/s Premium furniture	Chennai

Parts

pno	pname	part_spec
P1	Table	Wood
P2	Chair	Wood
P3	Table	Steel
P4	Almirah	Steel
P5	Almirah	Wood

Question Continues in Next Slide

The primary key of each table is indicated by underlining the constituent fields.

```
SELECT s.no, s.name
FROM Suppliers s, Catalogue c
WHERE s.sno = c.sno AND
      cost > (SELECT AVG (cost)
              FROM Catalogue
              WHERE pno = 'p4'
              GROUP BY pno);
```

225

The number of rows returned by the above SQL query is [MCQ: 2020-2M]

(A) 4

(B) 5

(C) 0

(D) 2

Q.

The following relation records the age of 500 employees of a company, where empNo (indicating the employee number) is the key: $empAge(\underline{empNo}, age$

Consider the following relational algebra expression:

$$\Pi_{empNo}(empAge \bowtie_{(age>age1)} \rho_{empNo1,\ age1(}empAge))$$

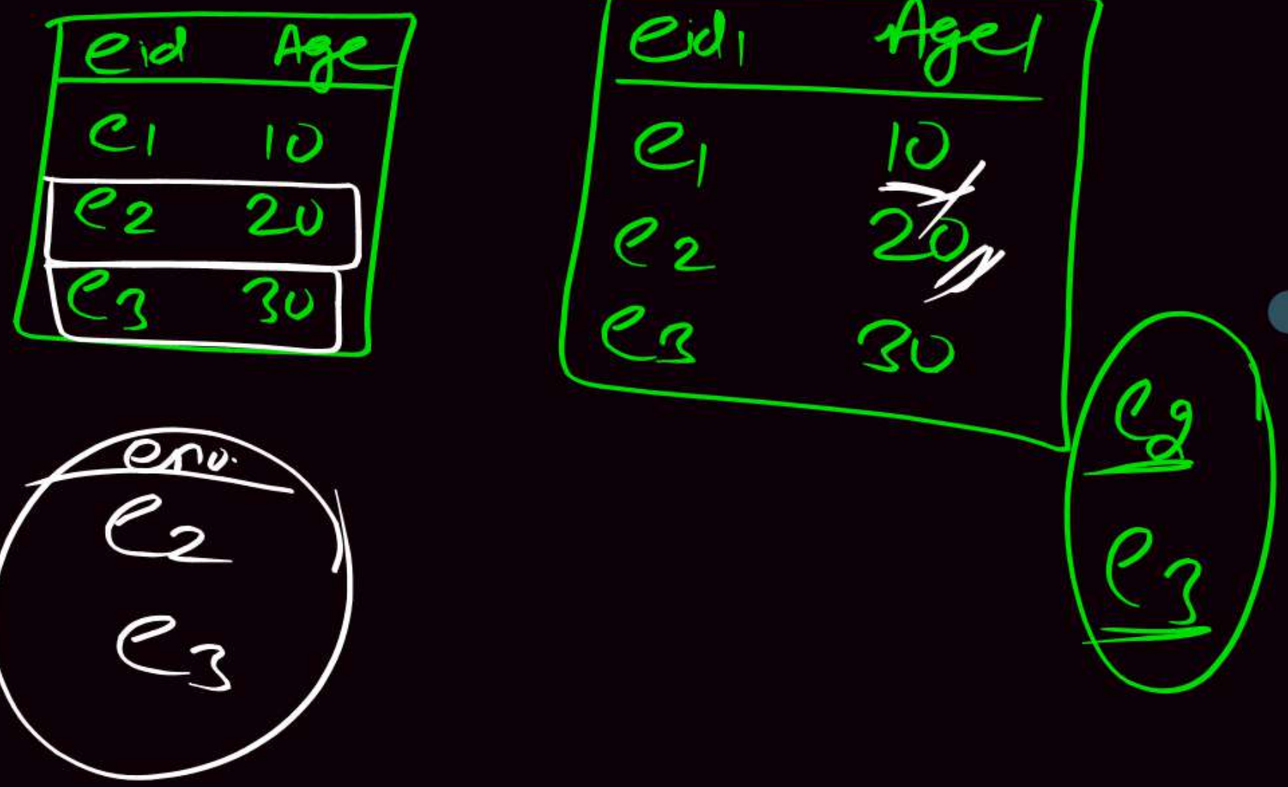
What does the above expression generate? [MCQ:2021-2M]

A. Employee numbers of only those employees whose age is the maximum.

B. Employee numbers of only those employees whose age is more than the age of exactly one other employee.

C. Employee numbers of all employees whose age is not the minimum.

D. Employee numbers of all employees whose age is the minimum.


eid Age
e1 10
e2 20
e3 30
eid1 Age1
e1 10
e2 20
e3 30
e2
e3
e1 10 > 10
10 > 20
10 > 30
e2 20 > 10
e2 20 > 20
e2 20 > 30
e3 30 > 10
e3 30 > 20
e3 30 > 30

Q.

Consider the relational database with the following four schemes and their respective instances.

Student(sNo, sName, dNo) Dept(dNo, dName)
Course(cNo, cName, dNo) Register(sNo, cNo)

Student

sNo	sName	dNo
S01	James	D01
S02	Rocky	D01
S03	Jackson	D02
S04	Jane	D01
S05	Milli	D02

Dept

dNo	dName
D01	CSE
D02	EEE

Course

cNo	cName	dNo
C11	DS	D01
C12	OS	D01
C21	DE	D02
C22	PT	D02
C23	CV	D03

Register

sNo	cNo
S01	C11
S01	C12
S02	C11
S03	C21
S03	C22
S03	C23
S04	C11
S04	C12
S05	C11
S05	C21

Question Continues in Next Slide

SQL Query:

```
SELECT * FROM Student AS S WHERE NOT EXIST
(SELECT cNo FROM Course WHERE dNo = "D01"
EXCEPT
SELECT cNo FROM Register WHERE sNo = S.sNo)
```

The number of rows returned by the above SQL query is ____.

[NAT: 2022: 2M]

S01: C11 C12 EXCEPT C11 C12 ⇒ Empty

S02: C11 C12 EXCEPT C11 ⇒ C12

S03: C11 C12 EXCEPT C21 C22 C23 ⇒ C11 C12

S04: C11 C12 EXCEPT C11 C12 ⇒ Empty

S05: C11 C12 EXCEPT C11 C21 ⇒ C12

EXCEPT

A, B
EXCEPT
B

A

A, B
EXCEPT
A

B

A, B
EXCEPT
A, B

Empty

A, B
EXCEPT
C, D

A, B

Relational Algebra and Tuple Calculus

Q.

Common Data for Next two question:

Consider the following relations A, B and C:

A.

ID	Name	Age
12	Arun	60
15	Shreya	24
99	Rohit	11
25	Hari	40
98	Rohit	20

B.

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

C.

ID	Phone	Age
10	2200	02
99	2100	01

[MCQ: 2012: 2M]

How many tuples does the result of the following relational algebra expression contain? Assume that the schema of $A \cup B$ is the same as that of A.

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

(A) 7 (B) 4 (C) 5 (D) 9

Q.

Consider two relations $R_1(A, B)$ with the tuples (1, 5), (3, 7) and $R_2(A, C) = (1, 7), (4, 9)$. Assume that R(A, B, C) is the full natural outer join of R_1 and R_2. Consider the following tuples of the form (A, B, C): a = (1, 5, null), b = (1, null, 7), c = (3, null, 9), d = (4, 7, null), e = (1, 5, 7), f = (3, 7, null), g = (4, null, 9).

Which one of the following statements is correct? [MCQ: 2015–1M]

A	B
1	5
3	7

A	C
1	7
4	9

A	B	C
1	5	7
3	7	NULL
4	NULL	9

A. R contains a, b, e, f, g but not c, d

B. R contains all a, b, c, d, e, f, g

C. R contains e, f, g but not a, b

D. R contains e but not f, g

Q.

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

The following query is made on the database.

- $T_1 \leftarrow \pi_{\text{CourseName}}(\sigma_{\text{StudentName} = \text{'SA'}}(\text{CR}))$
- $T_2 \leftarrow \text{CR} \div T_1$

The number of rows in T_2 is _____.

[NAT:2017–2M]

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

Q.

Consider the following relations P(X, Y, Z), Q(X, Y, T) and R(Y, V)

P

X	Y	Z
X1	Y1	Z1
X1	Y1	Z2
X2	Y2	Z2
X2	Y4	Z4

Q

X	Y	T
X2	Y1	2
X1	Y2	5
X1	Y1	6
X3	Y3	1

R

Y	V
Y1	V1
Y3	V2
Y2	V3
Y2	V2

How many tuples will be returned by the following relational algebra query? [NAT:2019-2M]

$$\Pi_x(\sigma_{(P.Y=R.Y \wedge R.V=V2)}(P \times R)) - \Pi_x(\sigma_{(Q.Y = R.Y \wedge Q.T > 2)}(Q \times R))$$

2 − 1 = 1

Transactions and Concurrency Control

Q.

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

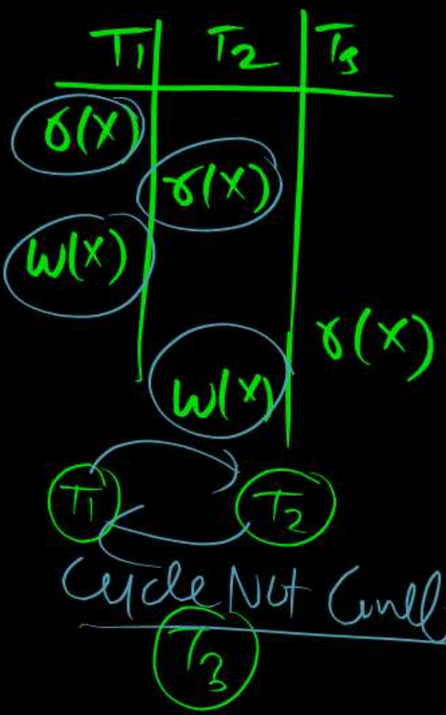
[MCQ:2014–2M]

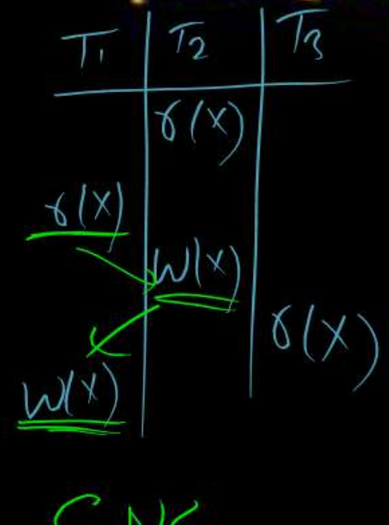
A $r_1(x); r_2(x); w_1(x); r_3(x); w_2(x)$

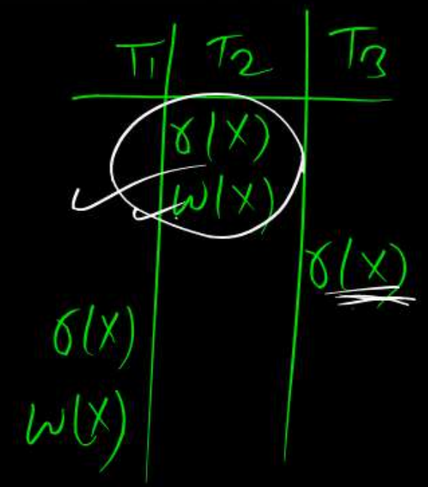
B $r_2(x); r_1(x); w_2(x); r_3(x); w_1(x)$

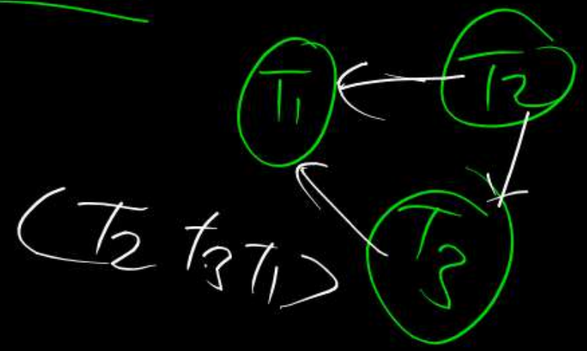
C $r_3(x); r_2(x); r_1(x); w_2(x); w_1(x)$

D $r_2(x); w_2(x); r_3(x); r_1(x); w_1(x)$









Conflict Operation

Same Data Item

R(A) – W(A)

W(A) – R(A)

W(A) – W(A)

Q. Consider the following transaction involving two bank accounts x and y.

read (x); x := x - 50; write (x); read (y); y:= y + 50; write (y)

The constraint that the sum of the accounts x and y should remain constant is that of

[MCQ:2015–1M]

A. Atomicity

B. Consistency

C. Isolation

D. Durability

Q.

Consider a simple checkpointing protocol and the following set of operations in the log.

(start, T4); (write, T4, y, 2, 3); (start, Tl);
(commit, T4); (write, T1, z, 5, 7);
(checkpoint);
(start, T2); (write, T2, x, 1, 9); (commit, T2);
(start, T3); (write, T3, z, 7, 2);

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

[MCQ: 2015–2M]

A. Undo: T3, T1; Redo: T2

B. Undo: T3, T1; Redo: T2, T4

C. Undo: none; Redo: T2, T4, T3, Tl

D. Undo: T3, Tl, T4; Redo: T2

— (Transaction DataItem OLDValue NewValue)

		OLDValue	NewValue
T_1	X	5	7
T_1	X	7	11

$\equiv$

X = 5 ≠ 11

Redo

X = 11

X = 11 ≠ 5

Undo: X = 5

Q.

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

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

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

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

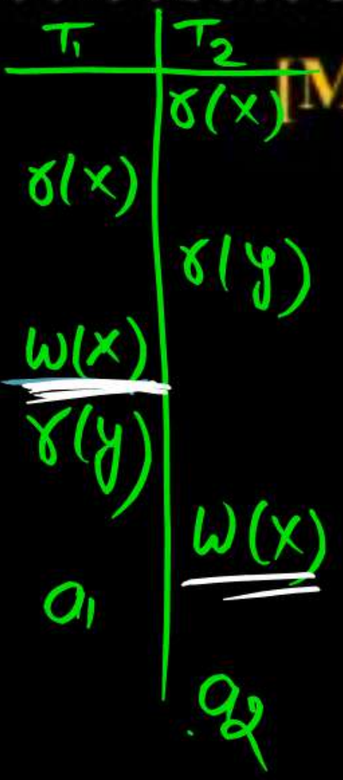
[MCQ:2016–2M]

A. S is non-recoverable

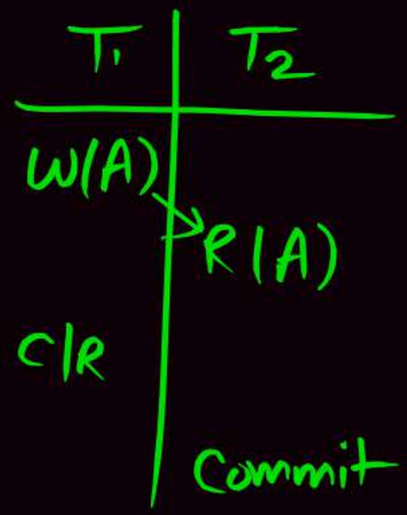
B. S is recoverable, but has a cascading abort

C. S does not have a cascading abort

D. S is strict

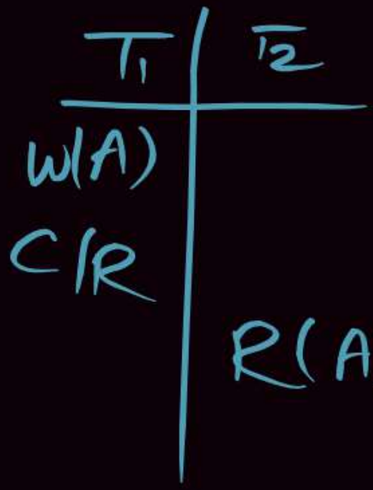


✓ Recoverable
✓ Cascadeless
✗ Strict Recoverable



Recoverable

Cascadeless Schedule



Strict Recoverable

Q.

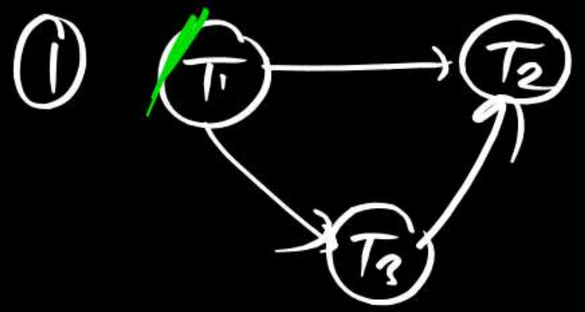
Suppose a database schedule S involves transactions $T_1 \ldots\ldots T_n$. Construct the precedence graph of S with vertices representing the transactions and edges representing the conflicts. If S is serializable, which one of the following orderings of the vertices of the precedence graph is guaranteed to yield a serial schedule? **[MCQ:2016–1M]**

(A) Topological Order

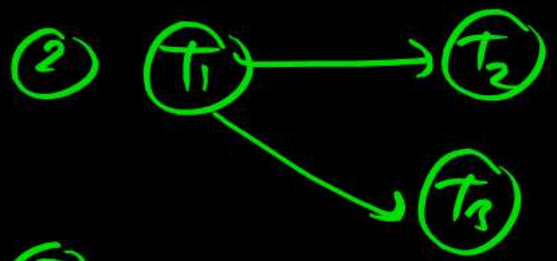
(B) Depth-first order

(C) Breadth-first order

(D) Ascending order of transaction indices.

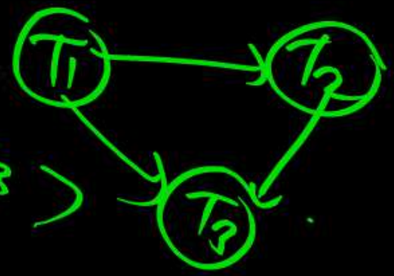


<T1 T3 T2>



<T1 T2 T3>
OR
<T1 T3 T2>

3



<T1 T2 T3>

Q. Which one of the following is NOT a part of the ACID properties of database transactions?

[MCQ:2016–1M]

A Atomicity

B Consistency

C Isolation

D Deadlock - freedom

Q.

Consider the following two statements about database transaction schedules:

I. Strict two-phase locking protocol generates conflict serializable schedules that are also recoverable.

II. Timestamp-ordering concurrency control protocol with Thomas' Write Rule can generate view serializable schedules that are not conflict serializable.

Which of the above statements is/are TRUE? [MCQ:2019–1M]

A. I only

B. II only

C. Both I and II

D. Neither I nor II

write

$TS(T_i) < RTS(Q)$

If $TS(T_i) < WTS(Q)$; No Rollback

2PL

Growing Phase
Shrinking Phase

Conflict Serializability

But Irrecoverable

STRICT 2PL

2PL + All X lock
Release After Commit / Rollback

Recoverable
Cascadeless
STRICT Recoverable

Q.

Consider a schedule of transactions T1 and T2;

T1	RA			RC		WD		WB	commit	
T2		RB	WB		RD		WC			commit

Here, RX stands for "Read(X)" and WX stands for "Write(X)". Which one of the following schedules is conflict equivalent to the above schedule?

[MCQ:2020-2M]

$R_2(B) - W_1(B)$
$W_2(B) - W_1(B)$
$R_1(C) - W_2(C)$
$R_2(D) - W_1(D)$

$R(B) - W(B)$: $T_2 \rightarrow T_1$

(A)

T1				RA	RC	WD	WB		Commit	
T2	RB	WB	RD					WC		Commit

(B)

T1	RA	RC	WD	WB					Commit	
T2					RB	WB	RD	WC		Commit

$R(B) - W(B)$
$T_2 \rightarrow T_1$

(C)

T1	RA	RC	WD				WB		Commit	
T2				RB	WB	RD		WC		Commit

(D)

T1					RA	RC	WD	WB	Commit	
T2	RB	WB	RD	WC						Commit

$R_2(B) - W_1(B)$
$W_2(B) - W_1(B)$
$R_1(C) - W_2(C)$
$R_2(D) - W_1(D)$

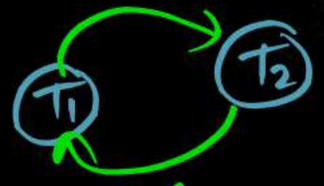
Q.

Let $r_i(z)$ and $w_i(z)$ denote read and write operations respectively on a data item by a transaction T_i. Consider the following two schedules.

S_1: $r_1(x)$ $r_1(y)$ $r_2(x)$ $r_2(y)$ $w_2(y)$ $w_1(x)$

S_2: $r_1(x)$ $r_2(x)$ $r_2(y)$ $w_2(y)$ $r_1(y)$ $w_1(x)$

Which one of the following options is correct? [MCQ: 2021: 2M]



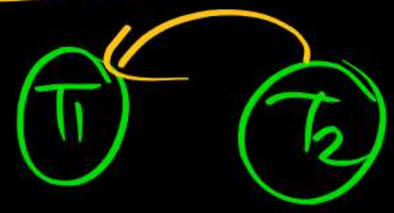
CNC
Cyclic Not Conflict

(A) S_1 is conflict serializable, and S_2 is not conflict serializable.

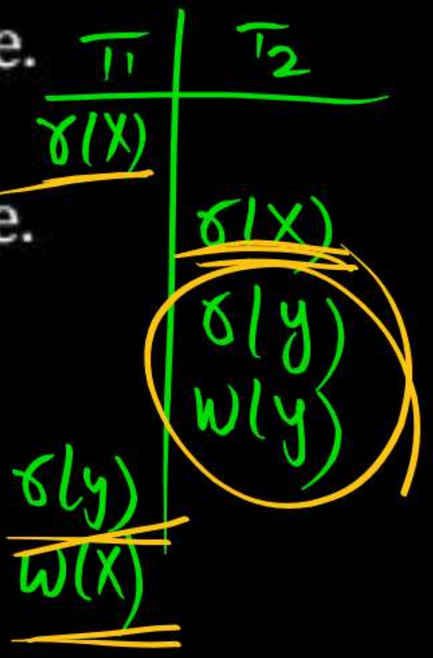
(B) S_1 is not conflict serializable, and S_2 is conflict serializable.

(C) Both S_1 and S_2 are conflict serializable.

(D) Neither S_1 nor S_2 is conflict serializable.



<T2 T1>
Conflict



Q.

Let S be the following schedule of operations of three transactions T_1, T_2 and T_3 in a relational database system:

$$R_2(Y), R_1(X), R_3(Z), R_1(Y), W_1(X), R_2(Z), W_2(Y), R_3(X), W_3(Z)$$

Consider the statements P and Q below:

P: S is conflict-serializable.

Q: If T_3 commits before T_1 finishes, then S is recoverable.

Which one of the following choices is correct?

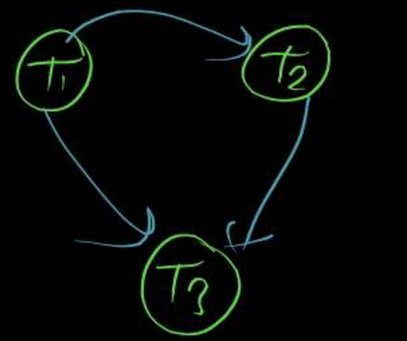
[MCQ: 2021-2M]

(A) Both P and Q are true.

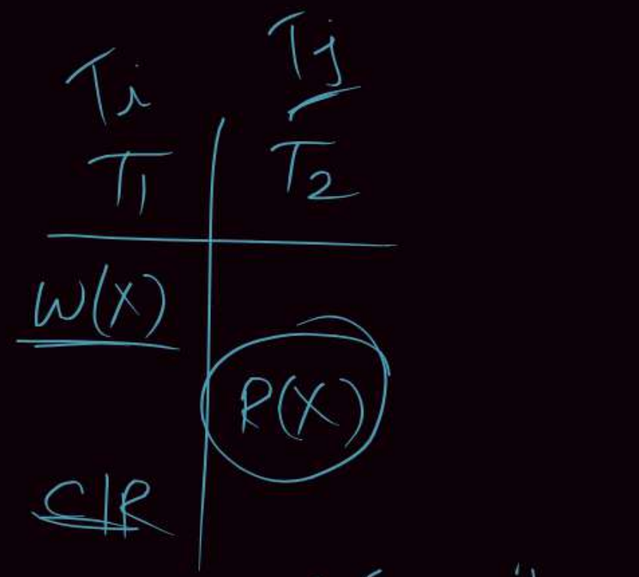
(B) P is true and Q is false.

(C) P is false and Q is true.

(D) Both P and Q are false.



$\langle T_1\ T_2\ T_3 \rangle$



Recoverable

File Structures

Q. 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?

[MCQ: 2010–1M]

A 1

B 2

C 3

D 4

Maximum keys = 5

ORDER[P] = 6

ORDER : P

keys = P-1

Min # keys = $\lceil \frac{P}{2} \rceil - 1$

$\Rightarrow \lceil \frac{6}{2} \rceil - 1$

$\Rightarrow 3-1 = 2$ Ans

Q.

Consider a B^+ tree in which the search key is 12 bytes long, block size is 1024 bytes, record pointer is 10 bytes long and block pointer is 8 bytes long. The maximum number of keys that can be accommodated in each non-leaf node of the tree is_____.

[NAT: 2015–2M]

Internal

keys = 12 Byte Block Size = 1024 Byte R_P = 10 Byte B_P = 8 Byte

$P * B_P + (P-1)\,key \leq Block\ Size$

$P*8 + (P-1)*12 \leq 1024$

$8P + 12P - 12 \leq 1024$

$20P \leq 1036$

$P = \frac{1036}{20} \Rightarrow \lfloor 51.5 \rfloor = 51$

key = P − 1

⇒ 51 − 1

⇒ 50 Ans

Q.50

B⁺ Trees are considered BALANCED because

[MCQ: 2016–1M]

A. The lengths of the paths from the root to all leaf nodes are all equal.

B. The lengths of the paths from the root to all leaf nodes differ from each other by at most 1.

C. The number of children of any two non-leaf sibling nodes differ by at most 1.

D. The number of records in any two leaf nodes differ by at most 1.

Tree

Binary Tree

BST

AVL

$b_f = 0, 1, -1$

LR RL RR LL

Q.

Which one of the following statements is NOT correct about the B^+ tree data structure used for creating an index of a relational database table?

[MCQ:2019–1M]

A. Key values in each node are kept in sorted order.

B. Each leaf node has a pointer to the next leaf node.

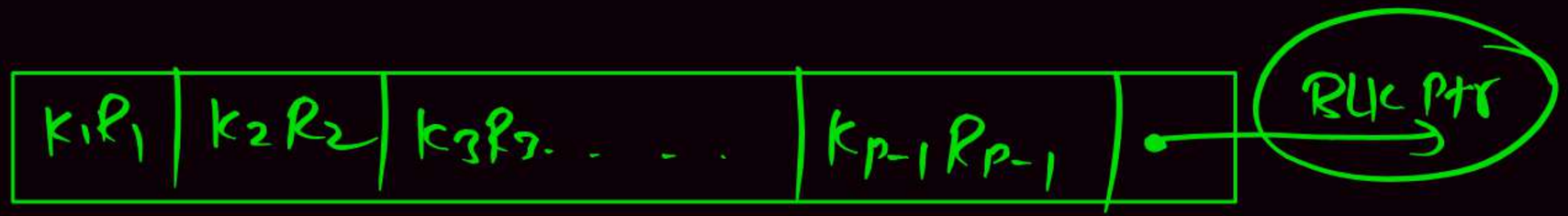
C. B^+ tree is a height balanced tree.

D. Non-leaf nodes have pointers to data records.

Internal

B+ Tree

Leaf Node



Q.

A data file consisting of 1,50,000 student-records is stored on a hard disk with block size of 4096 bytes. The data file is sorted on the primary key RollNo. The size of a record pointer for this disk is 7 bytes. Each student-record has a candidate key attribute called ANum of size 12 bytes. Suppose an index file with records consisting of two fields, ANum value and the record pointer to the corresponding student record, is built and stored on the same disk. Assume that the records of data file and index file are not split across disk blocks. The number of blocks in the index file is _____.

[NAT:2021-1M]

Block Size = 4096 Byte

key = 12B Bp = 7 Byte

One Index Record Size = Size of key + Size of Bp

12 + 7 = 19

Block factor of Index file = $\left\lfloor \frac{4096}{19} \right\rfloor$ = 215

Index Block = $\left\lceil \frac{1,50,000}{215} \right\rceil$ = 698 Ans

THANK YOU!