



COMPUTER SCIENCE

Database Management System

Query Language

Lecture_9



Vijay Agarwal sir



**TOPICS
TO BE
COVERED**

01

Queries

02

SQL Clauses



① Relational Algebra

② SQL

- all clauses
- Operators
- Aggregate function
- Query writing.

JOIN with Foreign key Concept

Foreign key : key points:

Note The value present in foreign key must be present in Primary key of Referenced Relation.

Note Foreign key may contain Duplicates & NULL values.

Table Contain Foreign key
which

Referencing Relation (Table).

JOIN with Foreign key Concept

Note Whenever Two Table (Relation) are Joined (Natural Join) with respect to Primary key & foreign key then Maximum Number of tuples in the Resulting Relation is equals to Number of tuples in the Referencing Relation.

JOIN with Foreign key Concept

Note Foreign key : key points:

The value present in foreign key must be present in Primary key of Referenced Relation.

Note Foreign key may contain Duplicates & NULL values.

Table contain Foreign key
which

Referencing Relation (Table).

JOIN with foreign key concept

(Q.L)

Referencing Reln
R(A B C)

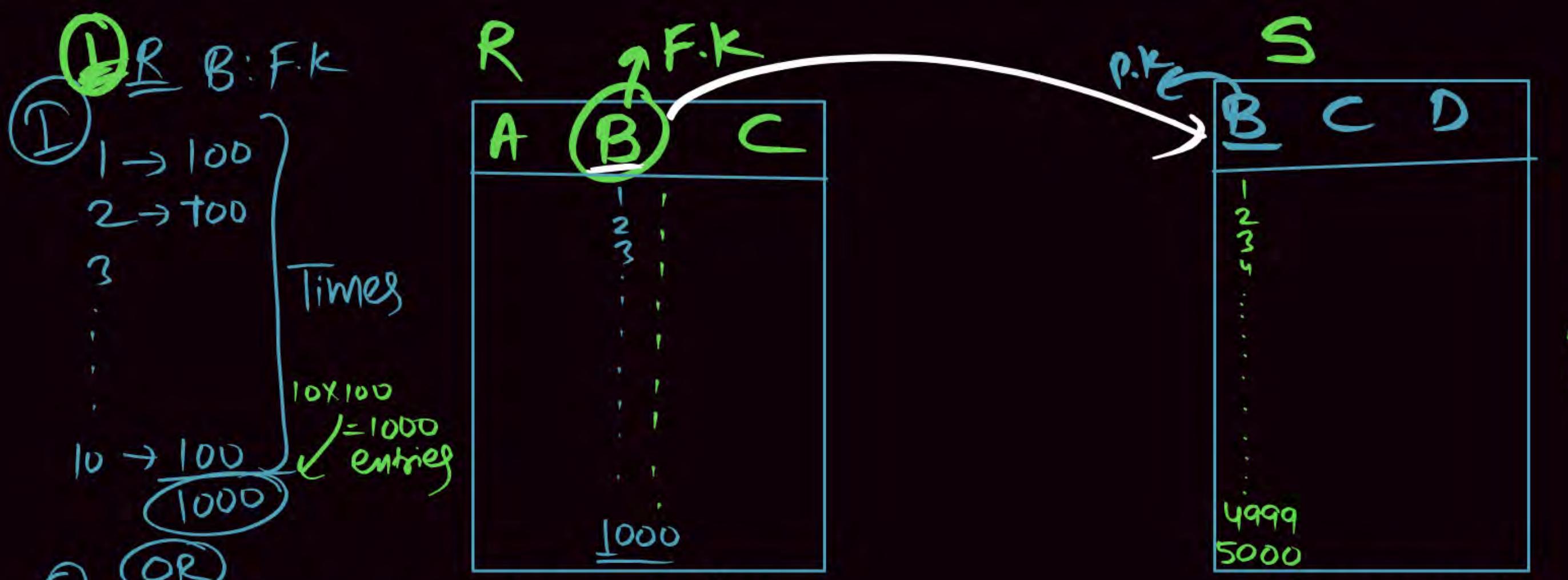
'B' is Foreign key.

- a) 1000
- b) 5000
- c) 5×10^6
- d) None of these.
Maximum Tuples in R \bowtie S = 1000 Ans

S(B C D)

Referenced Relation.

5000 Tuple



1000 Tables

5000 Tuples

Maximum

RMS = 1000

JOIN with foreign key Concept

(Q2)

Referencing Reln

R (A B C)

B is Foreign key.

1000 Tuple

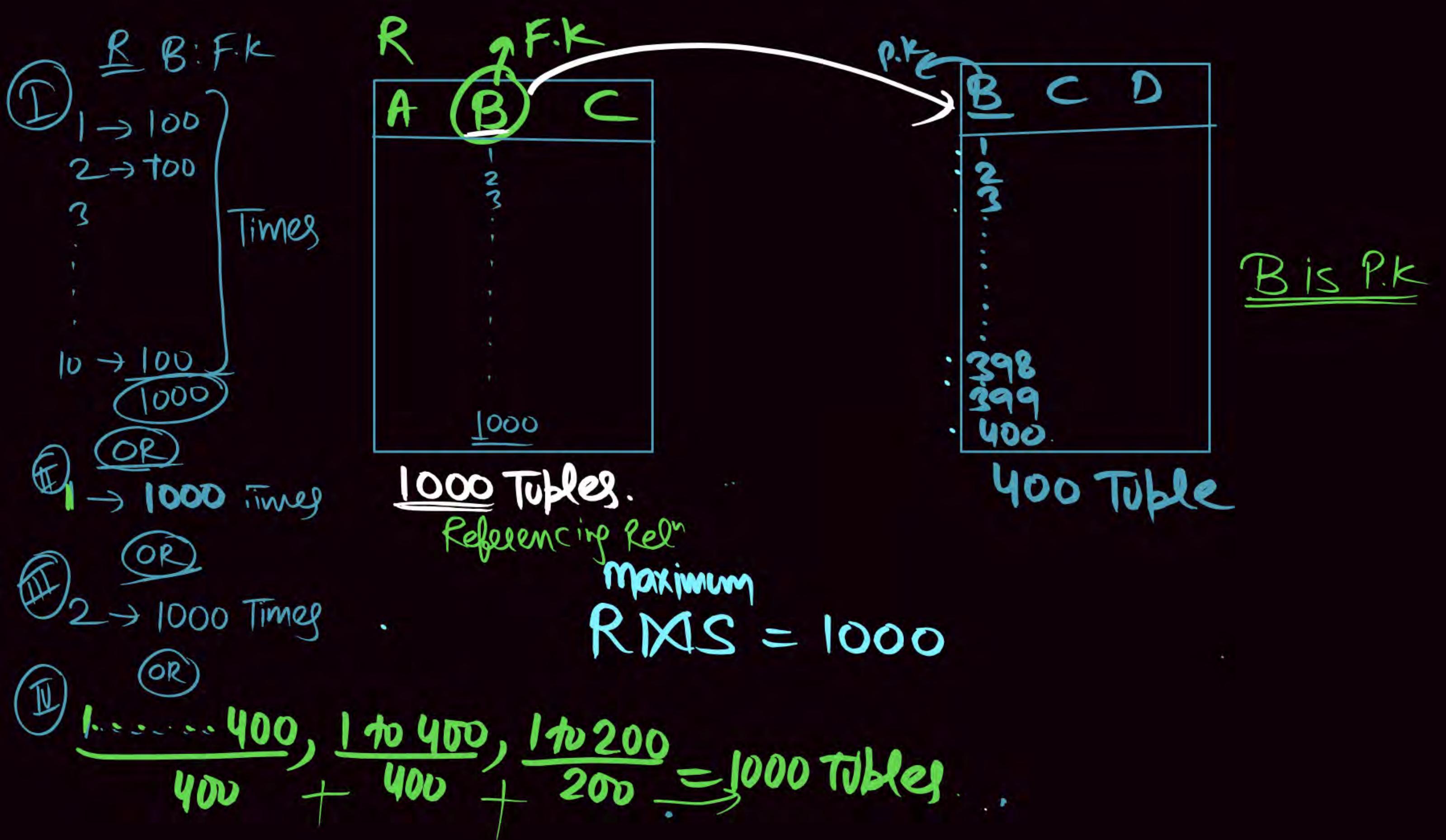
Referenced Reln

S (B C D)

400 Tuple

- (A) 400
- (B) 1000
- (C) 4×10^5
- (D) None of these.

Maximum Tuples in R \bowtie S = 1000 Avg



How To Find Minimum :

& its Variations

Q.1

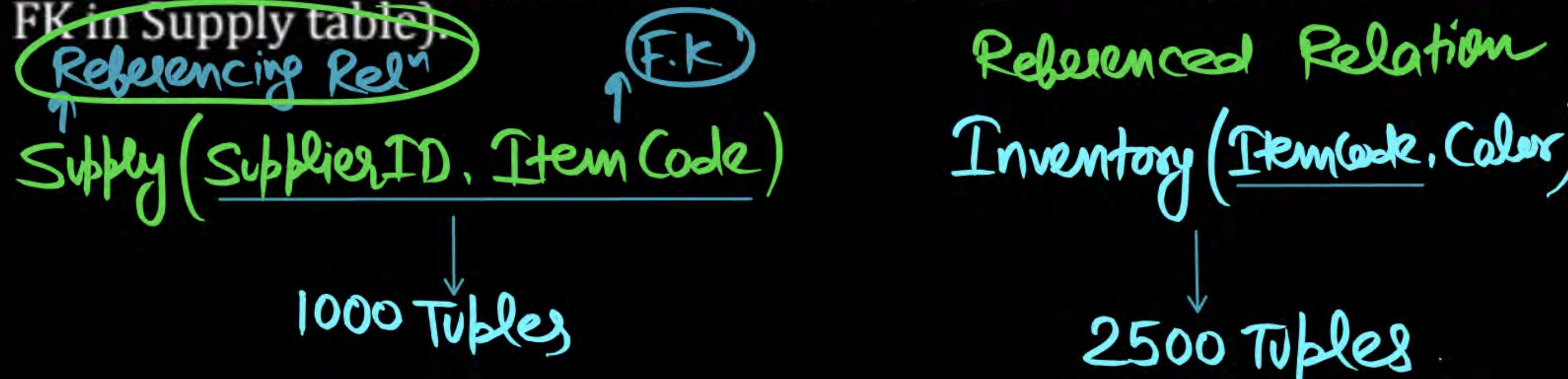
P
W

Consider the relations:

Supply (SupplierID, Itemcode) with 1000 tuples

Inventory (Itemcode, color) with 2500 tuples

Let p and q be the number of maximum and minimum records in Supply JOIN Inventory, the value of $p + q$ is 2000 Ans (Itemcode is FK in Supply table).



Maximum # Tuples = 1000

Minimum # Tuples = 1000

The Value Present in Foreign key Must be Present in Primary key.

Foreign key May Contain Duplicates & NULL Values.

Supply (Supplier id, Item Code)

Inventory (ItemCode, Color)

Item Code is Foreign key in Supply Table & (But)

Here Item Code is key also So Item Code Can not Not be NULL in this Case.

So Minimum Tuples = 1000

Supply

Item Code
Duplicate But
NOT NULL
(\because Item Code
is a key)

<u>SupplierID</u>	<u>ItemCode</u>
S ₁	P ₁
S ₁	P ₂
S ₁	P ₈
S ₂	P ₁
S ₂	P ₂
S ₄	P ₁

↑
Not NULL
(\because key)

Inventory

<u>ItemCode</u>	<u>Color</u>
P ₁	Red
P ₂	Green
P ₃	Yellow
P ₄	Red
.	.
.	.
P ₂₀₀₀	

Q. 2

P
W

Consider the relations:

Employee(Eid, Ename) with 2000 tuples

Department(Eid, Did) with 1000 tuples

Let p and q be the number of maximum and minimum records in Employee JOIN Department, the value of $p + q$ is 2000 (Eid is FK in Department table).

$$[p] \text{ Maximum} = 1000$$

$$[q] \text{ Minimum} = 1000$$

$$p+q = 2000$$

Q.3

Consider the relations:

$R_1(A, \underline{B}, C)$ with n tuples \rightarrow Referencing Relation

$R_2(\underline{B}, D, E)$ with m tuples (B is FK in R_1).

maximum tuples in $R_1 \bowtie R_2 = n$

minimum tuples in $R_1 \bowtie R_2$

n Tuples $\leftarrow R(A \underline{B} C)$
 B : F.K

$R_2(\underline{B} D E) \rightarrow m$ Tuples

Maximum # Tuples = n

Minimum # Tuples = n [If B is Foreign key & Not NULL]

= 0 [B is Not F.K, B is Not key @ B having NULL value]

Q.4

Consider the relations:

$R1(A, B, C)$ with n tuples

$R2(B, D, E)$ with m tuples (B is Not FK in R1).

maximum tuples in $R1 \bowtie R2$: n

minimum tuples in $R1 \bowtie R2$: 0

Q.5

R: m tuple

S: n tuple

Consider the join $R \bowtie S$ between R (A B C) and S (A D E) with attribute 'A' being the primary key in both relations and attribute A in S is a Foreign Key referencing to attribute A in R. R has m tuples and S has n tuples, maximum and minimum number of tuples in $R \bowtie S$ respectively are?

- (a) m, m + n
(b) m*n, 0
(c) min (m, n), min (m, n)
(d) max (m, n), min (m, n)

R	A	B	C
1.			
2.			
3.			
4.			
5.			
6.			

Referenced Relation

A: unique + Not Null P
W
& A is Primary key f
A is Foreign key

S	A	D	E
1.			
2.			
3.			
4.			
5.			
6.			

Referencing Relation.



$S(\underline{A} D E)$

A is Foreign key

& A is also Primary key

So Here the Value of A is Relation S

Contain NOT only NULL, But Not Duplicate algo.

A in S : Must be Unique + NOT NULL But those

Value of 'A' Must be Present in 'A' of Relation R (\because Foreign key Concept)

Q. 6

Consider the relations:

$R_1(A, B, C)$ with n tuples

$R_2(D, B, E)$ with m tuples

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

maximum tuples in $R_1 \bowtie R_2$

minimum tuples in $R_1 \bowtie R_2$

B is Non key Attribute in Both
Relation R & S.

Maximum = $n \times m$ Ans

Minimum = 0 Ans

$R_1(\underline{ABC})$

$(A)^+ = [ABC]$

A is P.K

R

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

$R_2(\underline{DBE})$ P W

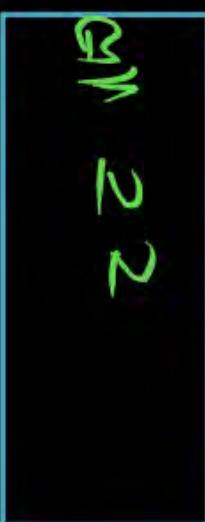
$(D)^+ = [DBE]$

D : Primary key.

S

D	B	E
2	1	
3	1	
4	1	
5	1	
6	1	
7	1	

D	B	E
7		
7		
7	8	
8	b	
b	b	



MCQ

The following functional dependencies hold for relations $R(A, \underline{B}, C)$ and $S(\underline{B}, D, E)$ FD for both the tables.

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

$$(B)^+ = \{BAC\} \quad B \text{ is Primary key of } R$$

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 $R \bowtie S$?

$R(A \underline{B} C)$ $S(\underline{B} D E)$
B: PK
Referenced Reln
Referencing Reln

[GATE-2010-CS: 2M]

A 100

200

100

B 200

C 300

D 2000

Q.

Consider two relations $R_1(A, B)$ with the tuples $(1, 5), (3, 7)$ and $R_2(A, C) = (1, 7)(4, 9)$

	A	B	C
a	1	5	N
b	1	N	7
c	3	N	9
d	4	7	NULL
e	1	5	7
f	3	7	NULL
g	4	N	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 = $(1, 5, \text{null})$, b = $(1, \text{null}, 7)$, c = $(3, \text{null}, 9)$, d = $(4, 7, \text{null})$, e = $(1, 5, 7)$, f = $(3, 7, \text{null})$, g = $(4, \text{null}, 9)$. Which one of the following statements is correct?

[GATE-2015: 1 Mark]

R_1	R_2
A	C
1 5	1 7
3 7	4 9

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

B R contains all of 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

$R \not\models S$
 $R \not\models S$

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

e
f
g

NAT

P
W

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?

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

$$X_2 - X_1 = X_2 = 1 \text{ Ans}$$

[GATE-2019-CS: 2M]

MCQ

Suppose $R_1(A, B)$ and $R_2(C, D)$ are two relation schemes. Let r_1 and r_2 be the corresponding relation instances. B is a foreign key that refers to C in R_2 . If data in r_1 and r_2 satisfy referential integrity constraints, which of the following is ALWAYS TRUE?

[GATE-2013-CS: 2M]

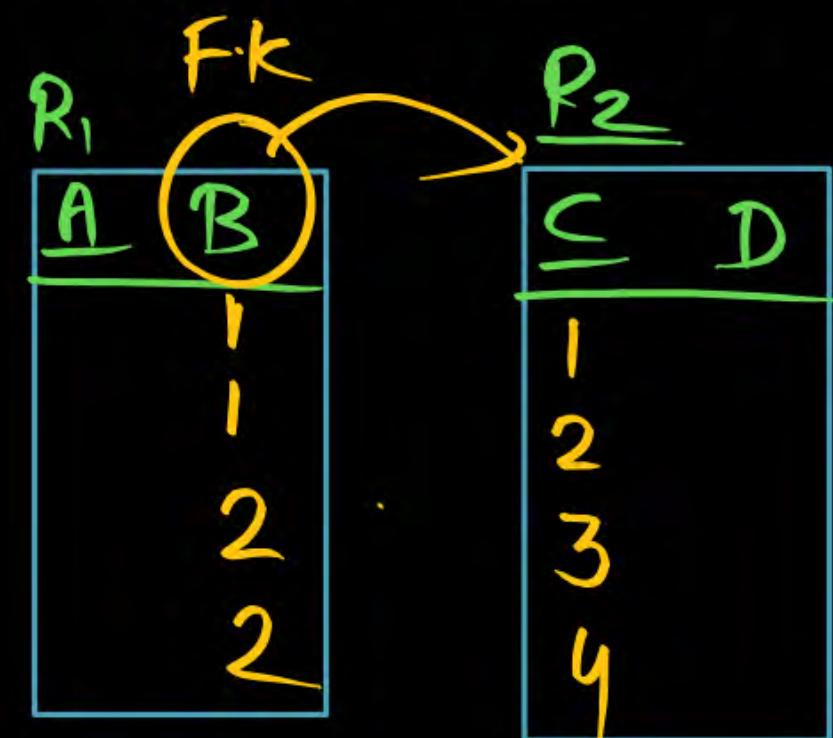
C [1,2] - (1,2,3,4) $\vdash \phi$

A $\Pi_B(r_1) \cdot \Pi_C(r_2) = \phi$

C $\Pi_B(r_1) = \Pi_C(r_2)$

B $\Pi_C(r_2) \cdot \Pi_B(r_1) = \phi$

D $\Pi_B(r_1) \cdot \Pi_C(r_2) \neq \phi$



Consider the following relation A, B and C:

A		
ID	Name	Age
12	Arun	60
15	Shreya	24
99	Rohit	11

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

C		
ID	Phone	Area
10	2200	02
99	2100	01

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.

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

[GATE-2012-CS: 2M]

A 7

B 4

C 5

D 9

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

P
W

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 ____.

[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

Q.

Consider the following relations:

P
W

Student	
Roll_No	Student_Name
1	Raj
2	Rohit
3	Raj

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

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.

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]

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 ____.

EC	Count
Dept Name	Num
AA	4
AB	3
AC	3
AD	2
AE	1

[NAT:2017-1M]

$$\frac{4+3+3+2+1}{5}$$

$$\frac{13}{5} = 2.6$$

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

Q.

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

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

Performance		
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 form 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 _____.

[NAT: 2019–2M]

WHEN we Do cross Product then

$$5 \times 5 = 25 \text{ Table}$$

But we do group By Student Name

then Only 5 Table in o/p.

5 Ans

Q.

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

C
C

```
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
```

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

[NAT:2016-2M]

Q Ans

Total

Name	Capacity
Ajmer	20
Bikaner	40
Churu	30
Dungarpur	10

Total-Avg.

Ang. Capacity
25

Name
Bikaner
Churu

Ans

Ans(2)

$$\frac{20+40+30+10}{4} = \frac{100}{4} = 25$$

Total Capacity > Total-Avg Capacity.

$$\left. \begin{matrix} 20 \\ 40 \\ 30 \\ 10 \end{matrix} \right\} > 25$$

Q.

Suppose $R_1(A, B)$ and $R_2(C, D)$ are two relation schemas. Let r_1 and r_2 be the corresponding relation instances. B is a foreign key that refers to C in R_2 . If data in r_1 and r_2 satisfy referential integrity constraints, which of the following is **ALWAYS TRUE?**

P
W**[MCQ: 2012–2M]**

- A $\Pi_B(r_1) - \Pi_C(r_2) = \phi$
- B $\Pi_C(r_2) - \Pi_B(r_1) = \phi$
- C $\Pi_B(r_1) = \Pi_C(r_2)$
- D $\Pi_B(r_1) - \Pi_C(r_2) \neq \phi$

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 field

```
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);
```

The number of rows returned by the above SQL query is

[MCQ: 2020-2M]

- A 4
- B 5
- C 0
- D 2

MCQ

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

Let MX and MY denote the respective maximum values of X and Y among all records in the table at any point in time. Using MX and MY, new records are inserted in the table 128 times with X and Y values being MX+1, 2*MY+1 respectively. It may be noted that each time after the insertion, values of MX and MY change.

What will be the output of the following SQL query after the steps mentioned above are carried out?

SELECT Y FROM T WHERE X=7;

[GATE-2011-CS: 2M]

A 127

B 255

C 129

D 257

(i) $x=2$,
 $y=2^2-1=3$

(ii) $x=3$
 $y=2^3-1=7$

(iii) $x=5$
 $y=2^5-1=31$

(iv) $x=7$
 $y=2^7-1=127$

$mx + l$

$$1+1=2$$

$$2+1=3$$

$$3+1=4$$

$$4+1=5$$

$$5+1=6$$

$$6+1=7$$

$$7+1=8$$

$$8+1=9$$

x	y
1	1
2	3
3	7
4	15
5	31
6	63
7	127
8	255
9	511
...	...

$2*my + l$

$$2*1+1=3$$

$$2*3+1=7$$

$$2*7+1=15$$

$$2*15+1=31$$

$$2*31+1=63$$

$$2*63+1=127$$

$$2*127+1=255$$

$$2*255+1=511$$

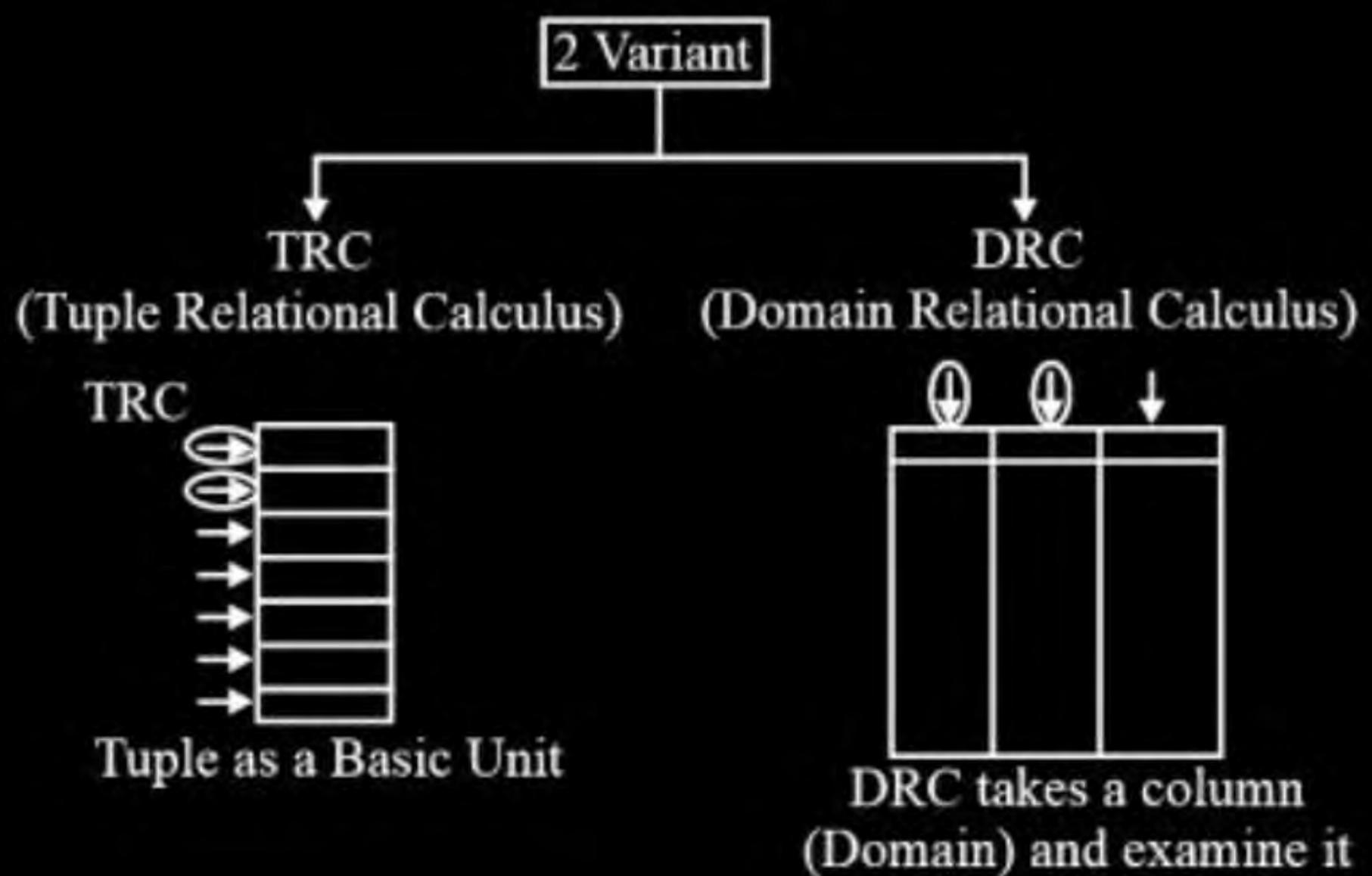
$$y = \boxed{2^x - 1}$$

Ars

Home work

Relational Calculus (Based on predicate calculus)

Non procedural query language and has



TRC

- Query describe result in the form of set of Tuples.

Tuple Variable

- It is variable that takes on Tuples of a relation schema as value.

DRC

- Query describe result in the form of set of Column. (Domain).

Domain Variable

- It is variable that Range over the domain of some attribute (column).

Form of Query

$$\{T \mid P(T)\}$$

↑
Tuple Variable
↑
Formula which
Describe Tuple
variable

TRC

$$\{T|P(T)\}$$

T : Tuple variable

P(T) : Formula over Tuple (T)

Such that P(T) is satisfied

Form of Query

$$\{\langle x_1 x_2 x_3 \dots x_n \rangle \mid P \langle x_1 x_2 x_3 \dots x_n \rangle\}$$

↓
Domain variable

↓
Formula which
describe Domain
variable

Form of Query

T{Selected Sid}

P(T){ FROM Student
WHERE Condition }

NOTE: Formula of TRC is expressed using first order logic
(Predicate logic).

First Order Logic :
Variable are 2 Types



Free variable

Bounded variable

Quantifiers

1. \forall : For all (ALL)
2. \exists : There exists (Any)

Bounded Variable: If tuple variable is preceded by the quantifier then it is Bounded variable.

Belong

1. $\forall \epsilon \text{Supplier}$
2. $\exists \epsilon \text{Supplier}$

Free variable: If tuple variable is not bounded by the quantifier then it is free variable.

NOTE: The result of TRC should be free variable.

Bounded Variable: If tuple variable is preceded by the quantifier then it is Bounded variable.

Belong

1. $\forall \epsilon \text{ Supplier}$
2. $\exists \epsilon \text{ Supplier}$

Free variable: If tuple variable is not bounded by the quantifier then it is free variable.

NOTE: The result of TRC should be free variable.

Ques.

LOAN (Loanno, Branch name, Amount)

Q. Find Loan no of amount above 5000.

TRC: { $T | T \in \text{Loan} \text{ (Total amount} > 50,000\text{)}}$ }

DRC: { $\langle \text{Loanno., branch name, amount} \rangle | \langle \text{Loan no, Branch, amount} \rangle \in \text{Loan} \text{ (amount} > 50000\text{) }$ }

Q. Retrieve Sid of the supplier who supplied some Red color parts.

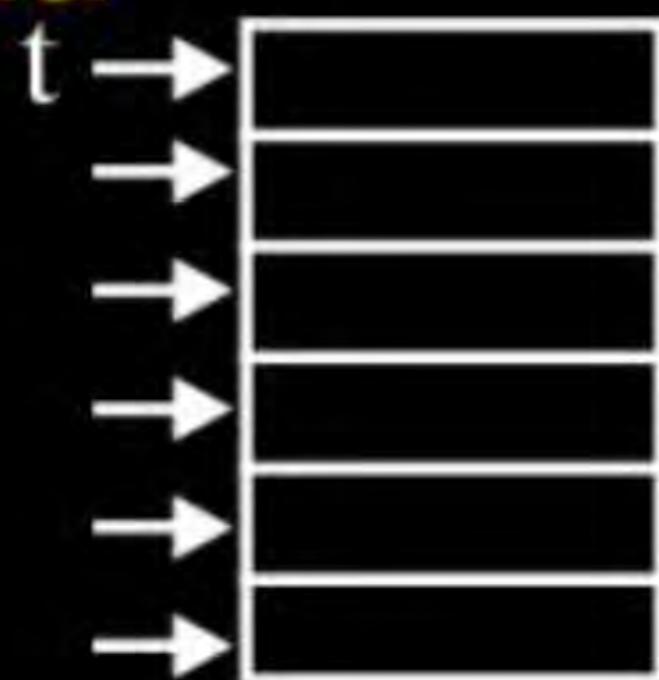
R.A:

$$\pi_{\text{Sid}} \left(\sigma_{\text{CPid} = \text{P.Pid}} \left(\text{Catalog} \times \sigma_{\text{colour} = \text{Red}} \left(\text{Parts} \right) \right) \right)$$

TRC:

$$(T | \exists C \in \text{Catalog} (\exists P \in \text{parts} (P.\text{color} = \text{Red} \wedge P.\text{Pid} = \text{CPid}) \wedge T = \text{CSid}))$$

TRC:

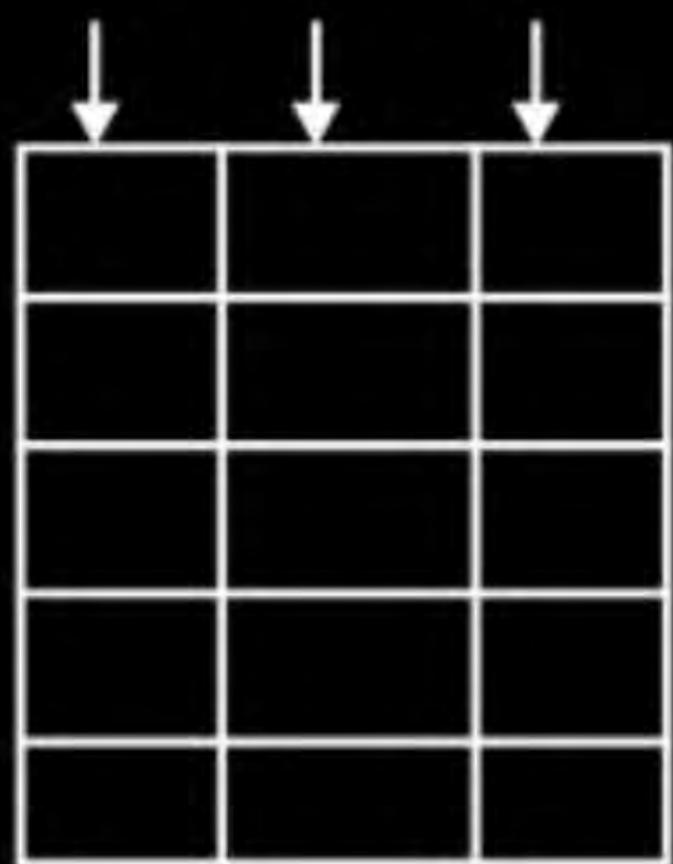


Takes a tuple and examine it

Tuple as a Basic unit

DRC: Takes a column

Range over the column or Domain and examine it



TRC: Unsafe operation occur

Unsafe operation: Student (t)

$\sim(t)$: Not belongs to student table (infinite/Universe all tuple)

So, its unsafe operation.

**THANK
YOU!**

