COMPUTER SCIENCE

Database Management System

Query Language

Lecture_2

Vijay Agarwal sir


TOPICS TO BE COVERED
01
Basic Operators
02
Derived Operators

Relational Algebra

Selection [σ] → Select the tuple $\sigma_{Condition}(R)$

Projection (Π) ⇒ $\Pi_{Attribute\ (or)\ Attribute\ List}(Relation\ R)$

- Union [$\cup$]
- Intersection [$\cap$]
- Minus | Set Difference [$-$]

Relational Algebra

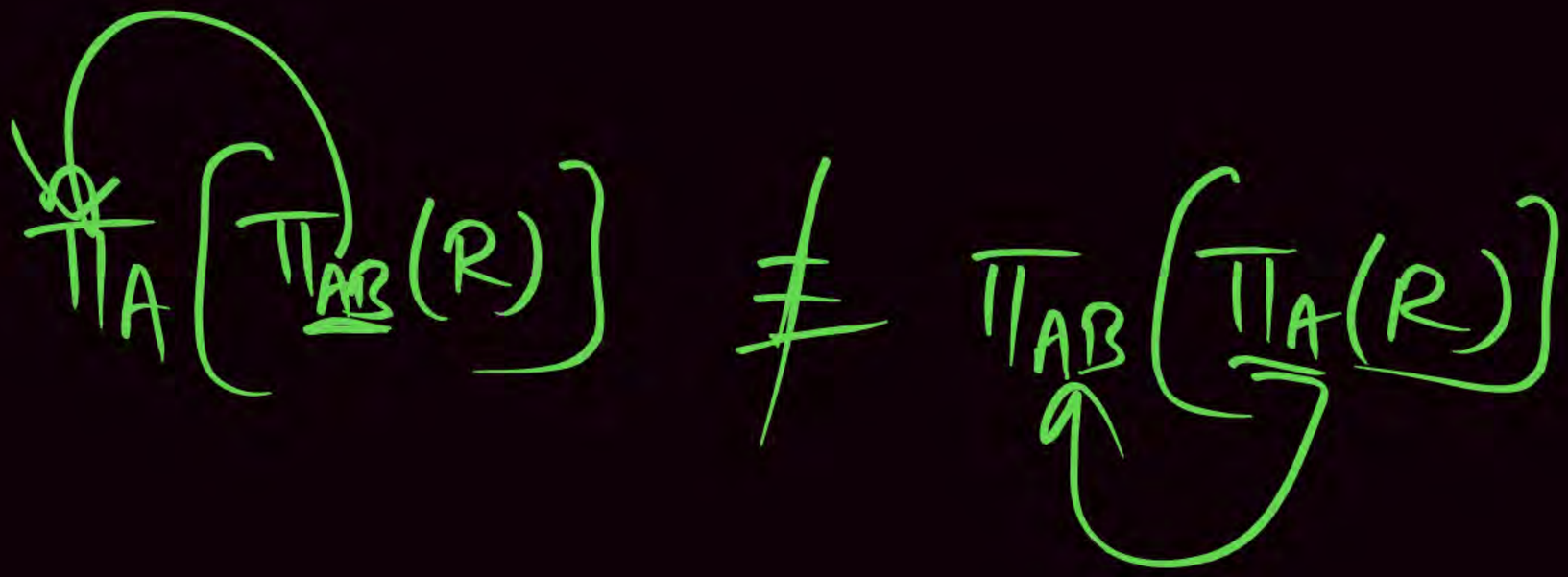
Basic operators

- π : Projection operator
- σ : Selection operator
- $\times$: Cross-product operator
- $\cup$: Union
- $-$: Set difference
- ϱ : Rename operator

Relational Algebra

Derived operators

- ∩ : Intersection {using "_"}
- ⋈ : Join {using X, σ}
- / or ÷ : Division {using π, x, −}


$\Pi_A(\Pi_{AB}(R)) \neq \Pi_{AB}(\Pi_A(R))$

CROSS PRODUCT [Cartesian Product]

R

n_1 Tuple

C_1 Attribute / Column

S

n_2 Tuple

C_2 Attribute / Column

$$R \times S = \frac{n_1 \times n_2 \text{ Tuple}}{C_1 + C_2 \text{ Attribute}}$$

Relational Algebra

Basic operators

II. Cross product (×):

- R × S : It result all attributes of R followed by all attributes of S, and each record of R paired with every record of S.

- Degree (R × S) = Degree (R) + Degree (S)
- |(R × S)| = |R| × |S|

C : Attribute
n : # Tuples

$C_1 + C_2$ Attribute

$n_1 \times n_2$ Tuple

NOTE:

- Relation R with n tuples and
- Relation S with 0 tuples then
- number of tuples in R × S = 0 tuples

CROSS PRODUCT[X]

R

A	B	C
1	2	4
3	2	6

S

B	C	D
2	4	8
2	7	4

R S

2 Tuple 2 Tuple

3 Attribute 3 Attribute

RXS = 2X2 = 4 Tuple
3+3=6 Attribute

R × S =

R.A	R.B	R.C	S.B	S.C	S.D
1	2	4	2	4	8
1	2	4	2	7	4
3	2	6	2	4	8
3	2	6	2	7	4

Q. WHY CROSS Product is Required ?

If We Require $R.C < S.D$?

(1) In Query language it is Not Possible to Fetch Directly because at a time It Fetch a Tuple either from Relation R (or) From Relation S. So Not Possible to Compare. That Why we Maintain RXS to Convert them into 1 Table So Comparsion possible.

(2) To Perform JOIN operation.

CROSS PRODUCT[X]

R

A	B	C
1	2	4
3	2	6

S

B	C	D
2	4	8
2	7	4

$R \times S =$

R.A	R.B	R.C	S.B	S.C	S.D
1	2	4	2	4	8
1	2	4	2	7	4
3	2	6	2	4	8
3	2	6	2	7	4

Cross – Product

Reserves (R_1)

Sid	bid	day
22	101	10/10/96
58	103	11/12/96

Sailors (S_1)

Sid	Sname	Rating	age
22	dustin	7	45.0
31	lubber	8	55.5
58	rusty	10	35.0

R_1 : 2 Tuple
3 Attribute

S_1 : 3 Tuple
4 Attribute

$R_1 \times S_1 = 2 \times 3 = 6$ Tuple
$3 + 4 = 7$ Attribute

S_1 R_1

(Sid)	Sname	Rating	age	(Sid)	bid	day
22	dustin	7	45.0	22	101	10/10/96
22	dustin	7	45.0	58	103	11/12/96
31	lubber	8	55.5	22	101	10/10/96
31	lubber	8	55.5	58	103	11/12/96
58	rusty	10	35.0	22	101	10/10/96
58	rusty	10	35.0	58	103	11/12/96

$S_1 \times R_1$

$R \bowtie S$

JOIN ($\bowtie$)

Natural Join ($\bowtie$)

Derived operator. Performed in 3 steps.

1. CROSS Product of R & S.
2. Select the tuples which satisfy equality condition on All Common Attributes (FROM R×S)
3. Projection of Distinct Attributes.

$$R \bowtie S = \pi_{\text{Distinct Attribute}}\left[\sigma_{\text{equality Condition on ALL common Attribute}}(R \times S)\right]$$

Step 3 → π (Distinct Attribute)

Step 2 → σ (equality Condition on ALL common Attribute)

Step 1 → $R \times S$

(Q) $R(A\,B\,\underline{C}\,\underline{D})$ $S(\underline{C}\,\underline{D}\,E\,F\,G)$

$$R \bowtie S = \pi_{ABCDEFG}\left[\sigma_{R.C = S.C \;\wedge\; R.D = S.D}(R \times S)\right]$$

$\wedge$ → AND

Join Operations

(1) Conditional Join ($\bowtie_c$)

(2) Equi join

(3) Natural join

(4) Left outer join

(5) Right outer join

(6) Full outer join

Conditional Join

$$R \bowtie_c S = \pi_{\text{ALL Attributes}}\left[\sigma_{\text{Given Condition}}\left(R \times S\right)\right]$$

~~Conditional~~ Natural Join

$R \bowtie_c S = \sigma_C (R \times S)$

$\Pi_{sid\,sname\,Rating\,age\,bid\,day}$

$\sigma_{R_1.sid = S_1.sid}(R \times S_1)$

S_1 Sid

R_1 Sid

$S_1 \times R_1$

~~$R \times S$~~

(Sid)	Sname	Rating	age	(Sid)	bid	day
22	dustin	7	45.0	22	101	10/10/96
22	dustin	7	45.0	58	103	11/12/96
31	lubber	8	55.5	22	101	10/10/96
31	lubber	8	55.5	58	103	11/12/96
58	rusty	10	35.0	22	101	10/10/96
58	rusty	10	35.0	58	103	11/12/96

Natural Join

$R_1 \bowtie S_1$:

Sid	Sname	Rating	age	bid	day
22	dustin	7	45	101	10/10/96
58	rusty	10	35	103	11/12/96

Conditional Join

$$R \bowtie_c S = \sigma_C (R \times S)$$

S_1 Sid

R_1 Sid

(Sid)	Sname	Rating	age	(Sid)	bid	day
22	dustin	7	45.0	22	101	10/10/96
22	dustin	7	45.0	58	103	11/12/96
31	lubber	8	55.5	22	101	10/10/96
31	lubber	8	55.5	58	103	11/12/96
58	rusty	10	35.0	22	101	10/10/96
58	rusty	10	35.0	58	103	11/12/96

$S_1 \times R_1$

~~$R_1 \times S_1$~~

Given Condition

$R_1.Sid > S_1.Sid.$

Conditional Join

$$R \bowtie_c S = \sigma_C (R \times S)$$

S_1 Sid

$S_1.Sid < R_1.Sid$

OR

$R_1.Sid > S_1.Sid$

R_1 Sid

$S_1 \times R_1$

$R \times S$

(Sid)	Sname	Rating	age	(Sid)	bid	day
22	dustin	7	45.0	22	101	10/10/96
22	dustin	7	45.0	58	103	11/12/96
31	lubber	8	55.5	22	101	10/10/96
31	lubber	8	55.5	58	103	11/12/96
58	rusty	10	35.0	22	101	10/10/96
58	rusty	10	35.0	58	103	11/12/96

Conditional Join

$S1 \bowtie_{S1.sid < R1.sid} R1$

output

(Sid)	Sname	Rating	age	(Sid)	bid	day
22	dustin	7	45.0	58	103	11/12/96
31	lubber	8	55.5	58	103	11/12/96

$S1 \bowtie_{S1.sid < R1.sid} R1$

- Result schema same as that of cross-product.
- Fewer tuples than cross – product, might be able to compute more efficiently.
- Sometimes called a theta –join.

Equi –Join

A special case of condition join where the condition c contains only equalities.

R_1 S_1

$R_1 \times S_1$

(Sid)	Sname	Rating	age	(Sid)	bid	day
22	dustin	7	45.0	22	101	10/10/96
22	dustin	7	45.0	58	103	11/12/96
31	lubber	8	55.5	22	101	10/10/96
31	lubber	8	55.5	58	103	11/12/96
58	rusty	10	35.0	22	101	10/10/96
58	rusty	10	35.0	58	103	11/12/96

Equi – Join

$S1 \bowtie_{sid} R1$

equality Join on Sid

Sid	Sname	Rating	age	bid	day
22	dustin	7	45.0	101	10/10/96
58	rusty	10	35.0	103	11/12/96

- Result schema similar to cross – product, but only one copy of fields for which equality is specified.
- Natural join: Equijoin on all common fields.

Equi Join

R(ABCD) S(CDEFG)

$$\Pi_{ABCDDEFG}\left[\sigma_{R.C=S.C}(R \times S)\right]$$

Natural Join

R(ABCD) S(CDEFG)

$$\Pi_{ABCDEFG}\left[\sigma_{\substack{R.C=S.C \wedge \\ R.D=S.D}}(R \times S)\right]$$

Join (⋈)

I. Natural join (⋈)

$R \bowtie S \equiv \pi_{\text{distinct attributes}}(\sigma_{\text{equality between common attributes of R and S}}(R \times S))$

Example:

- T_1 (ABC) and T_2 (BCDE)

$$\therefore \quad T_1 \bowtie T_2 = \pi_{ABCDE}\left(\sigma_{\substack{T_1 \cdot B = T_2 \cdot B \\ \wedge\, T_1 \cdot C = T_2 \cdot C}}(T_1 \times T_2)\right)$$

Note

- T_1 (AB) and T_2 (CD)

$$\therefore \quad T_1 \bowtie T_2 \equiv T_1 \times T_2 = \pi_{ABCD}(T_1 \times T_2)$$

NOTE:

Natural join equal to cross-product if join condition is empty.

Join (⋈)

II. Conditional Join ($\bowtie_c$)

- $R \bowtie_c S \equiv \sigma_c (R \times S)$

Join (⋈)

III. Outer Joins:

(a) LEFT OUTER JOIN

$R ⟕ S$: It produces

$(R ⋈ S) \cup$ {Records of R those are failed join condition with remaining attributes null}

(b) RIGHT OUTER JOIN (⟖)

$R ⟖ S$: It produces

$(R ⋈ S) \cup$ {Records of S those are failed join condition with remaining attributes null}

(C) FULL OUTER JOIN (⟗)

$R ⟗ S = (R ⟕ S) \cup (R ⟖ S)$

Natural Join ⋈

R

A	B	C
1	2	4
3	2	6

S

B	C	D
2	4	8
2	7	4

$\Pi_{ABCD}\left[\sigma_{R.B=S.B \wedge R.C=S.C}(R \times S)\right]$

$R \times S =$

R.A	R.B	R.C	S.B	S.C	S.D
1	2	4	2	4	8
1	2	4	2	7	4
3	2	6	2	4	8
3	2	6	2	7	4

$R \bowtie S =$

A	B	C	D
1	2	4	8

$$R \bowtie S = \pi_{ABCD} \left\{ \sigma_{\substack{R.B = S.B \\ R.C = S.C}} \Lambda^{(R \times S)} \right\}$$

$R \bowtie S =$

A	B	C	D
1	2	4	8

Left Outer Join $R \bowtie S$: $R \bowtie S$ & Tuples from Left Side Relation R which failed in Join Condition

Right Outer Join ($R \bowtie S$): $R \bowtie S$ & Tuples from Right Side Relation S, which Failed in Join Condition.

Left Outer Join [⟕]

$(R \bowtie S)$

R

A	B	C
1	2	4
3	2	6

S

B	C	D
2	4	8
2	7	4

Failed to satisfy Join condition

$(R \bowtie S) =$

A	B	C	D
1	2	4	8

$R \bowtie S =$

A	B	C	D
1	2	4	8
3	2	6	Null

$R \bowtie S =$

A	B	C	D
1	2	4	8
NULL	2	7	4

Right Outer Join [⟖]

$R \bowtie_C S =$

A	B	C	D
1	2	4	8
Null	2	7	4

$R \overline{\bowtie} S : R \overline{\bowtie} S \cup R \bowtie S$

Full Outer Join [⟗]

Full outer join = Left outer join Union Right outer join

$R ⟗ S = R ⟕ S \cup R ⟖ S$

R ⟕ S

A	B	C	D
1	2	4	8
3	2	6	Null

∪

R ⟖ S

A	B	C	D
1	2	4	8
Null	2	7	4

$R ⟗ S =$

A	B	C	D
1	2	4	8
3	2	6	Null
Null	2	7	4

Q.

Let R and S be two relations with the following schema

R(P, Q R1, R2, R3)

S(P, Q S1, S2)

R.P = S.P
R.Q = S.Q

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

[GATE : 2M
NIC - 2020]

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

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

III. $\pi_P(\pi_{P,Q} (R) \cap \pi_{P,Q}(S))$

IV. $\pi_P(\pi_{P,Q}(R) - (\pi_{P,Q} (R) - \pi_{P,Q} (S)))$

A Only I and II

B Only I and III

C Only I, II and III

D Only I, III and IV

R

P	Q	R_1	R_2	R_3
1	3			
2	4			
3	5			
4	6			
1	5			

S

P	Q	S_1	S_2
1	3		
2	1		
3	5		
4	5		

② $\Pi_P(R) \bowtie \Pi_P(S)$ →

P
1
2
3
4

③ $\Pi_P(\Pi_{P,Q} \cap \Pi_{P,Q})$

Π_P

P	Q
1	3
3	5

→

P
1
3

① $\Pi_P(R \bowtie S)$

$\Pi_P(R.P = S.P \wedge R.Q = S.Q\ (R \times S))$

P	Q
1	3
3	5

⇒ Π_P

1
3

Ans

④ is same as ③

$R \cap S = R - (R - S)$

R

P	Q	R_1	R_2	R_3
1	3			
2	4			
3	5			
4	6			

S

P	Q	S_1	S_2
1	7		
2	1		
3	11		
4	5		

② $\Pi_P(R) \bowtie \Pi_P(S)$ → P: 1, 2, 3, 4 ✓

③ $\Pi_P\left[\Pi_{P,Q} \cap \Pi_{P,Q}\right]$

① $\Pi_P(R \bowtie S)$

$\Pi_P\left[\begin{array}{l} R.P = S.P \wedge \\ R.Q = S.Q \end{array} (R \times S)\right]$

④ is same as ③ (→ ①)

$R \cap S = R - (R - S)$

Rename operator (ϱ)

It is used to rename table name and attribute names for query processing.

Example:

(I) Stud (Sid, Sname, age)

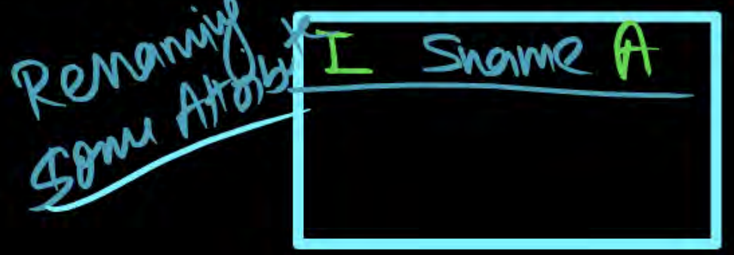
ϱ(Temp, Stud) : Temp (Sid, Sname, age)

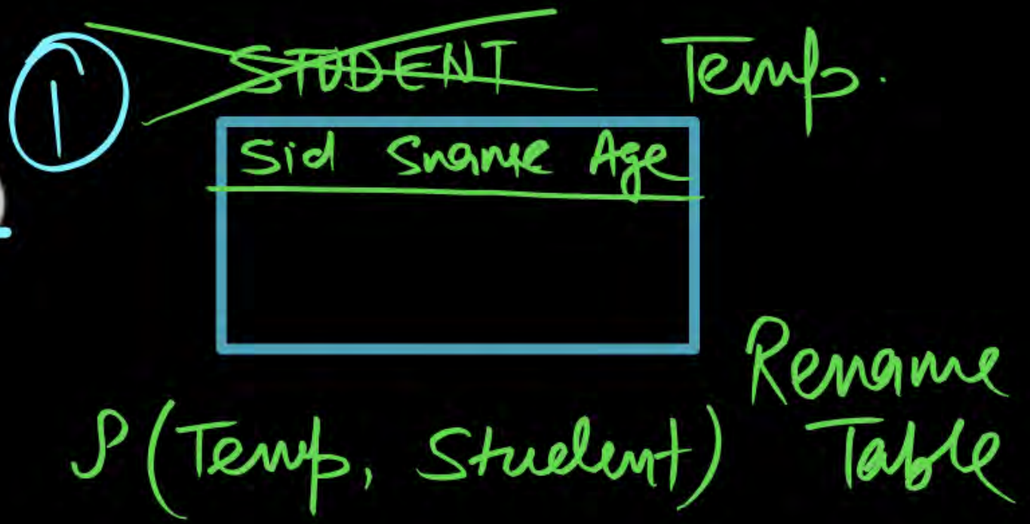
(II) $\varrho_{I, N, A}$ (Stud) : Stud (I, N, A)

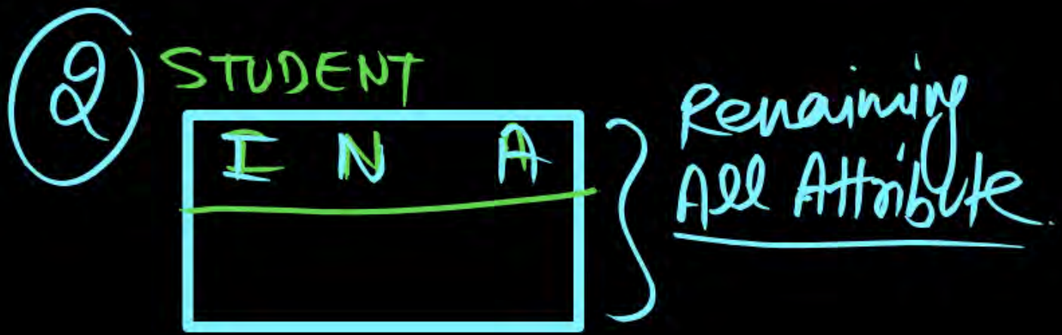
All attributes renaming

(III) $\varrho_{\substack{Sid \to I \\ age \to A}}$ (Stud) : Stud (I, Sname, A)

Some attribute renaming







Division

- It is used to retrieve attribute value of R which has paired with every attribute value of other relation S.
- $\pi_{AB}(R)/\pi_{B}(S)$: It will retrieve values of attribute 'A' from R for which there must be pairing 'B' value for every 'B' of S.

Expansion of '/' by using basic operator

- Example: Retrieve sid's who enrolled every course.
- Result:

  $\pi_{sidcid}(Enroll) / \pi_{cid}(Course)$

  Step 1: Sid's not enrolled every course of course relation.

  (Sid's enrolled proper subset of course)

  $\pi_{sid}((\pi_{sid}(Enroll) \times \pi_{cid}(course)) - \pi_{sidcid}(Enroll))$
- Step 2:

  [sid's enrolled every course] = [sid's enrolled some course] – [sid's not enrolled every course]

$\therefore \quad \pi_{sidcid}(E) / \pi_{cid}(c) = \pi_{sid}(E) - \pi_{sid}((\pi_{sid}(E) \times \pi_{cid}(C) - \pi_{sidcid}(E))$

Division

Q. Retrieve all student who are Enrolled **Some course** or **Any course** or at least one course?

Solution Π_{Sid} (Enrolled)

Enrolled	
Sid	**Cid**
S_1	C_1
S_1	C_2
S_1	C_3
S_2	C_1
S_2	C_3
S_3	C_1

Course
Cid
C_1
C_2
C_3

Division

Q. Retrieve all student who are Enrolled every course?

Solution

$\Pi_{Sid.Cid}$ (Enrolled) / Π_{Cid} (Course)

Find

2^{nd} attribute must be same.

Enrolled	
Sid	**Cid**
S_1	C_1
S_1	C_2
S_1	C_3
S_2	C_1
S_2	C_3
S_3	C_1

Course
Cid
C_1
C_2
C_3

Division

$$\Pi_{Sid}(\text{Enrolled}) - \Pi_{Sid}[\Pi_{Sid}(\text{Enrolled}) \times \Pi_{Cid}(\text{Course}) - \text{Enrolled}]$$

Division

$$\Pi_{AB}(R) \,/\, \Pi_{B}(S) = \Pi_{A}(R) - \Pi_{A}\,[\Pi_{A}(R) \times \Pi_{B}(S) - R\,]$$

Find

Connection

$$\Pi_{ABCD}(R) \,/\, \Pi_{CD}(S) \Rightarrow \Pi_{AB}(R) - \Pi_{AB}\,[\Pi_{AB}(R) \times \Pi_{CD}(S) - R\,]$$

Q.

Consider the following three relations in a relational database:

Employee (eId, Name), Brand (bId, bName), Own(eId, bId)

Which of the following relational algebra expressions return the set of eIds who own all the brands? [GATE: 2022]

A $\pi_{eId}\,(\pi_{eId,\,bId}\,(Own / \pi_{bId}\,(Brand))$

B $\pi_{eId}\,(Own) - \pi_{eId}\,((\pi_{eId}\,(Own) \times \pi_{bId}\,(Brand)) - \pi_{eId,\,bId}\,(Own))$

C $\pi_{eId}\,(\pi_{eId,\,bId}\,(Own) / \pi_{bId}\,(Own))$

D $\pi_{eId}\,((\pi_{eId}\,(Own) \times \pi_{bId}\,(Own) / \pi_{bId}\,(Brand))$

Consider the two relation Suppliers and Parts are given below.

Suppliers

S_{no}	P_{no}
S_1	P_1
S_1	P_2
S_1	P_3
S_1	P_4
S_2	P_1
S_2	P_2
S_3	P_2
S_4	P_2
S_4	P_4

Parts

P_{no}
P_2
P_4

$\pi_{S_{no}\,P_{no}}$ (Suppliers) / $\pi_{P_{no}}$ (Parts)

The number of tuples are there in the result when the above relational algebra query executes is _____.

Consider the Database with relations:

S Supplier (Sid, Sname, Rating)

P Parts (Pid, Pname, Color)

S Catalog (Sid Pid, Cost)

Q. Find the Sid of Supplier whose Rating greater than 9?

Q.

Find the Pid of Red Color Parts?

Q.

Retrieve Sid of Supplier whose cost is greater than 20,000?

Q.

Retrieve Sid of Supplier who supplied some Red color parts?

Solution: $\Pi_{Sid}\left[\sigma_{\substack{P.Pid=C.Pid \\ P.Color=Red}} \wedge (Catalog \times Parts)\right]$

Note: Let an Attribute A belongs to R only then

$\sigma_{A = 'a'} (R \bowtie S) = \sigma_{A = 'a'} (R) \bowtie S \rightarrow$ More efficiency query

Note: Let an Attribute A belongs to R only and Attribute B belongs to S only then

$\sigma_{A = 'a' \wedge B = 'b'} (R \bowtie S) = \sigma_{A = 'a'} (R) \bowtie \sigma_{B = 'b'} (S)$

Q.

Consider the following relation schemas:

b-Schema = (b-name, b-city, assets)

a-Schema = (a-num, b-name, bal)

d-Schema = (c-name, a-number)

Let branch, account and depositor be respectively instances of the above schemas. Assume that account and depositor relations are much bigger than the branch relation.

Consider the following query:

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

Q.

Which one of the following queries is the most efficient version of the above query? [GATE-2007: 2 Marks]

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

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

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

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

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? [GATE-2015: 1 Mark]

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

Summary

OPERATION	PURPOSE	NOTATION
SELECT	Selects all tuples that satisfy the selection condition from a relation R.	$\sigma_{\text{<selection condition>}}(R)$
PROJECT	Produces a new relation with only some of the attributes of R, and removes duplicate tuples.	$\pi_{\text{<attribute list>}}(R)$
THETA JOIN	Produces all combinations of tuples from R_1 and R_2 that satisfy the join condition.	$R_1 \bowtie_{\text{<join condition>}} R_2$
EQUIJOIN	Produces all the combinations of tuples from R_1 and R_2 that satisfy a join condition with only equality comparisons.	$R_1 \bowtie_{\text{<join condition>}} R_2$, OR $R_1 \bowtie_{\text{(<join condition 1 >), (<join condition 1 >)}} R_2$
NATURAL JOIN	Same as EQUIJOIN except that the join attributes of R_2 are not included in the resulting relation; if the join attributes have the same names, they do not have to be specified at all.	$R_1 *_{\text{<join condition>}} R_2$, OR $R_1 *_{\text{(<join attributes 1 >), (<join attributes 2 >)}} R_2$ OR $R_1 * R_2$

OPERATION	PURPOSE	NOTATION
UNION	Produces a relation that includes all the tuples in R_1 or R_2 or both R_1 and R_2; R_1 and R_2 must be union compatible.	$R_1 \cup R_2$
INTERSECTION	Produces a relation that includes all the tuples in both R_1 and R_2; R_1 and R_2 must be union compatible.	$R_1 \cap R_2$
DIFFERENCE	Produces a relation that includes all the tuples in R_1 and that are not in R_2; R_1 and R_2 must be union compatible.	$R_1 - R_2$
CARTESIAN PRODUCT	Produces a relation that has the attributes of R_1 and R_2 and includes as tuples all possible combinations of tuples from R_1 and R_2.	$R_1 \times R_2$
DIVISION	Produces a relation R(X) that includes all tuples t[X] in $R_1(Z)$ that appear in R_1 in combination with every tuple from $R_2(Y)$, where $Z = X \cup Y$	$R_1(Z) \div R_2(Y)$

THANK YOU!