

CSE 4125: Distributed Database Systems Chapter – 5

Translation of Global Queries to
Fragment Queries.
(Part – C)

Topics to be discussed –

- Equivalent Expression of Queries
- Two criteria for simplification of operator tree

Equivalent Expressions of Queries

- ❑ **Q1:** $PJ_{NAME, DEPTNUM} \bowtie SL_{DEPTNUM = 15} EMP$
- ❑ **Q2:** $SL_{DEPTNUM = 15} \bowtie PJ_{NAME, DEPTNUM} EMP$

$$Q1 \leftrightarrow Q2$$

Simplification of Operator Tree

Criterion – 1:

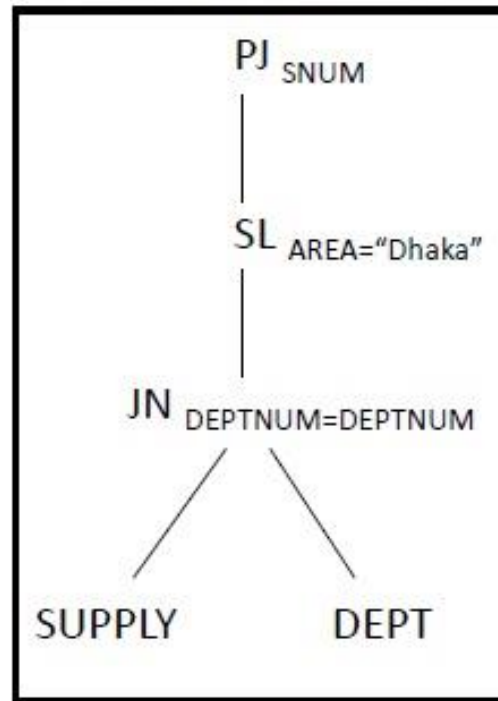
Appropriate introduce of *SL* and *PJ* in the tree.

- To get rid of unnecessary attributes.

SUPPLY (*SNUM*, *PNUM*, *DEPTNUM*, *QUAN*)
DEPT (*DEPTNUM*, *NAME*, *AREA*, *MGRNUM*)

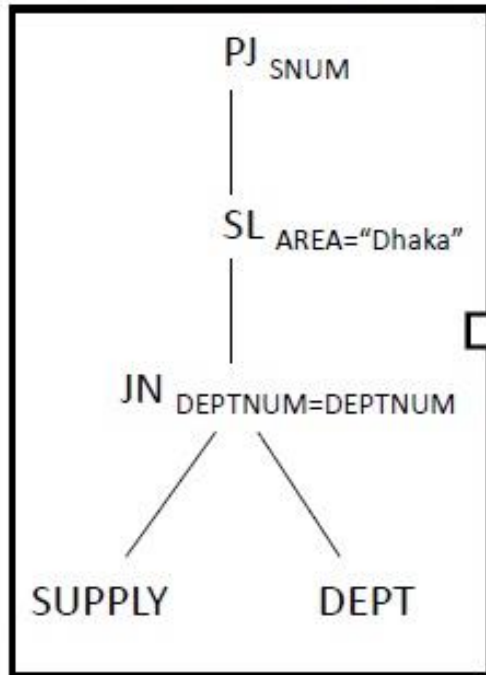
Q1: **PJ**_{SNUM} **SL**_{AREA="Dhaka"} (**SUPPLY** **JN**_{DEPTNUM=DEPTNUM} *DEPT*)

Operator Tree for Q1

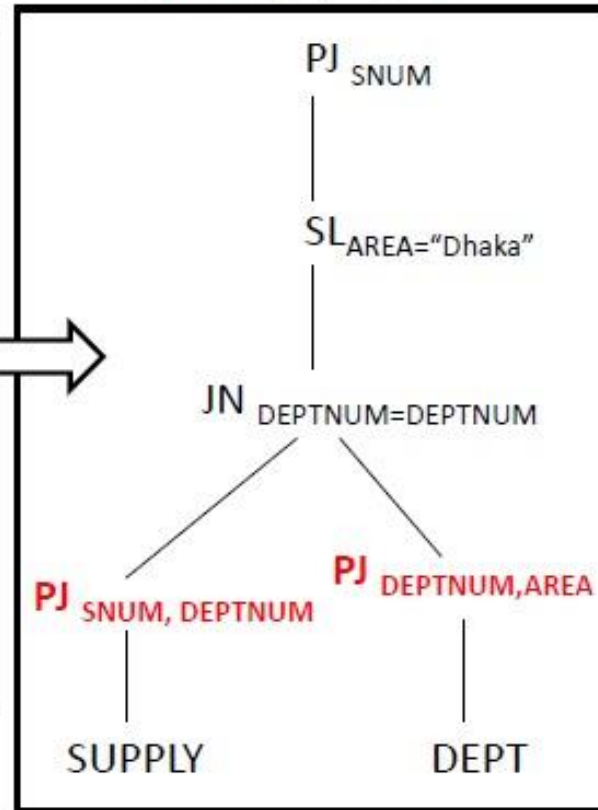


SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Operator Tree for Q1

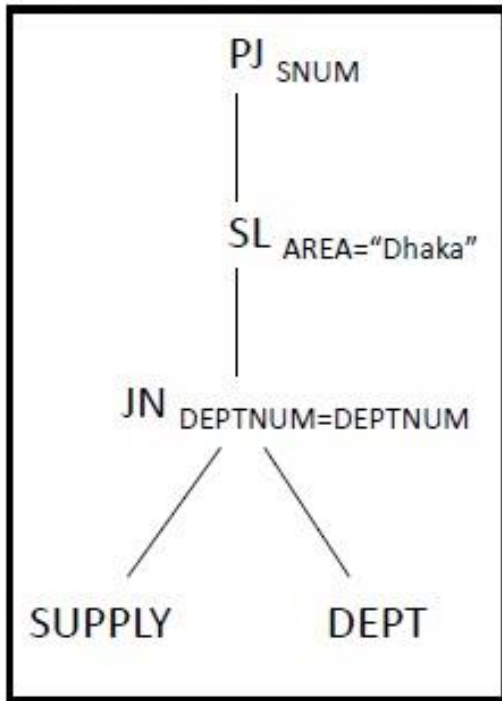


After applying Cr-1



SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Operator Tree for Q1



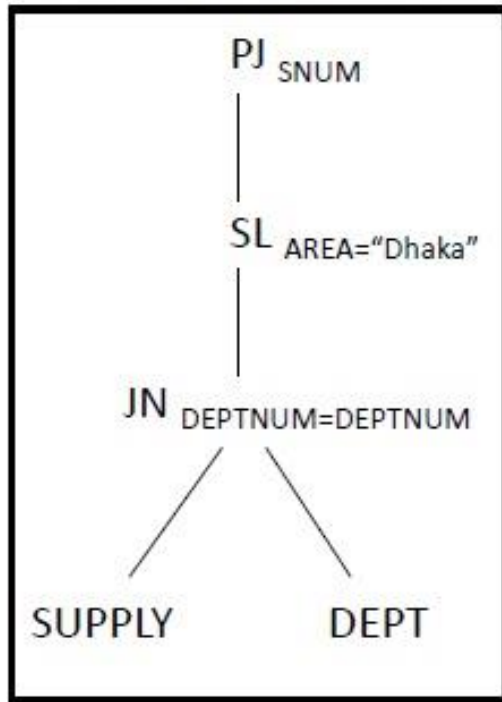
8 columns, 7 rows !

SNUM	PNUM	S.DEPTNUM	QUAN	D.DEPTNUM	NAME	AREA	MGRNUM
11		1		1		Dhaka	
12		2		2		Sylhet	
13		3		3		Rajshahi	
41		4		4		Dhaka	
51		5		5		Sylhet	
61		6		6		Rajshahi	
71		7		7		Dhaka	

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)

DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Operator Tree for Q1

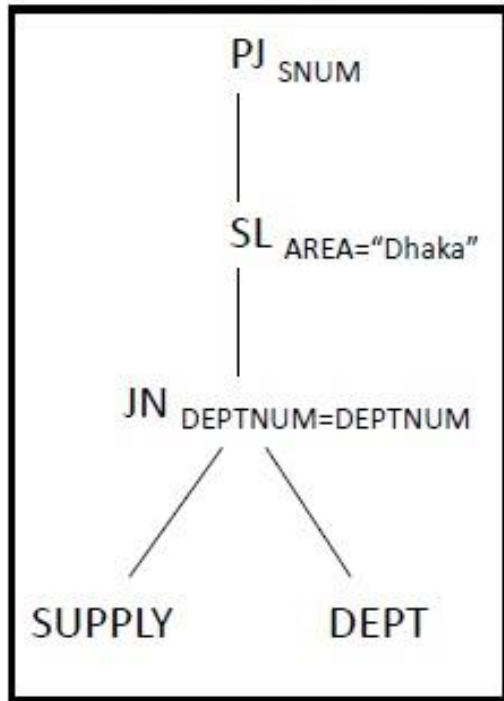


8 columns, 3 rows !

SNUM	PNUM	S.DEPTNUM	QUAN	D.DEPTNUM	NAME	AREA	MGRNUM
11		1		1		Dhaka	
41		4		4		Dhaka	
71		7		7		Dhaka	

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Operator Tree for Q1

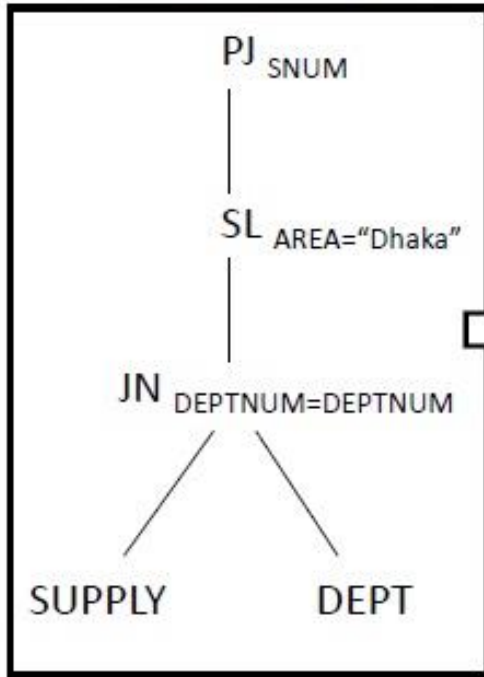


1 column, 3 rows !

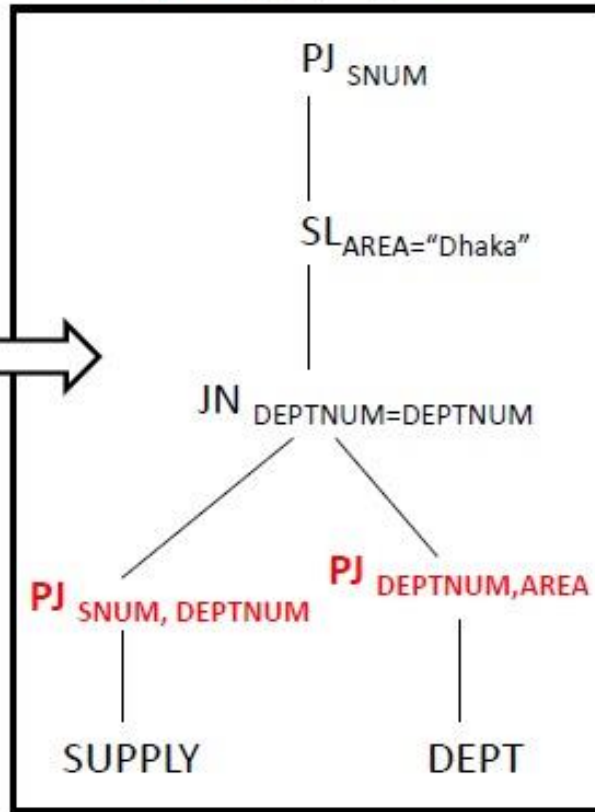
SNUM
11
41
71

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Operator Tree for Q1



After applying Cr-1



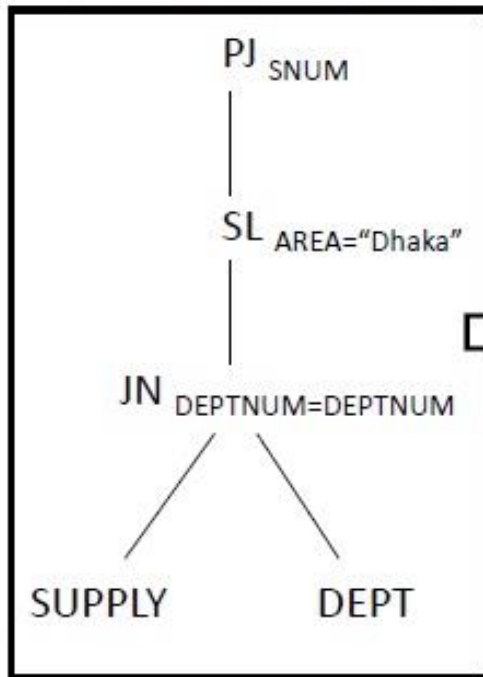
4 columns, 7 rows !

SNUM	S.DEPTNUM	D.DEPTNUM	AREA
11	1	1	Dhaka
12	2	2	Sylhet
13	3	3	Rajshahi
41	4	4	Dhaka
51	5	5	Sylhet
61	6	6	Rajshahi
71	7	7	Dhaka

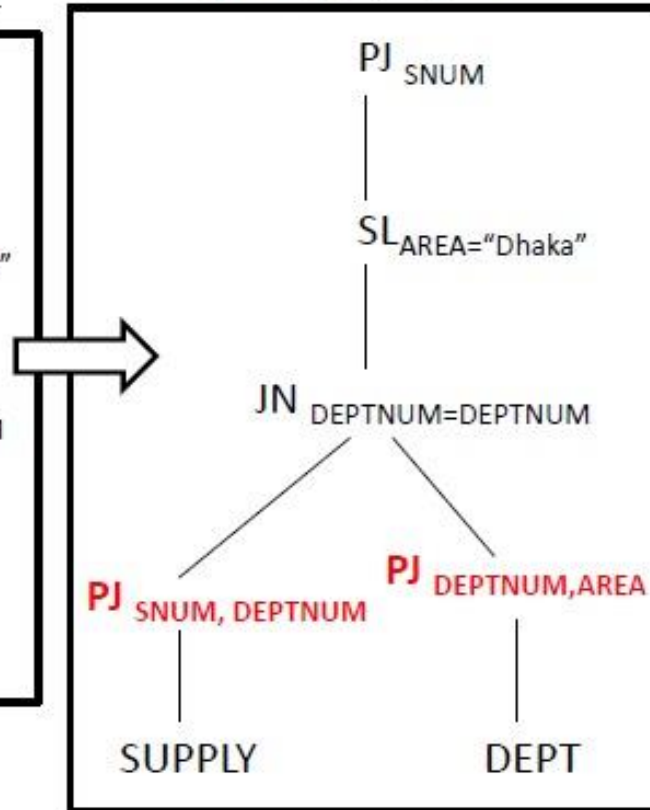
SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)

DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Operator Tree for Q1



After applying Cr-1

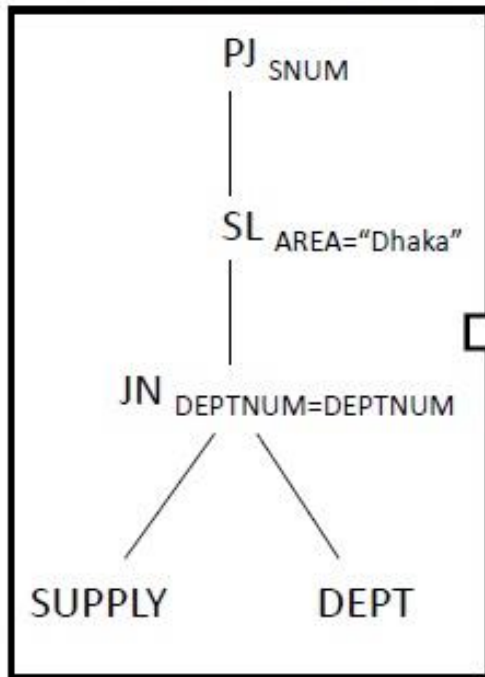


4 columns, 3 rows !

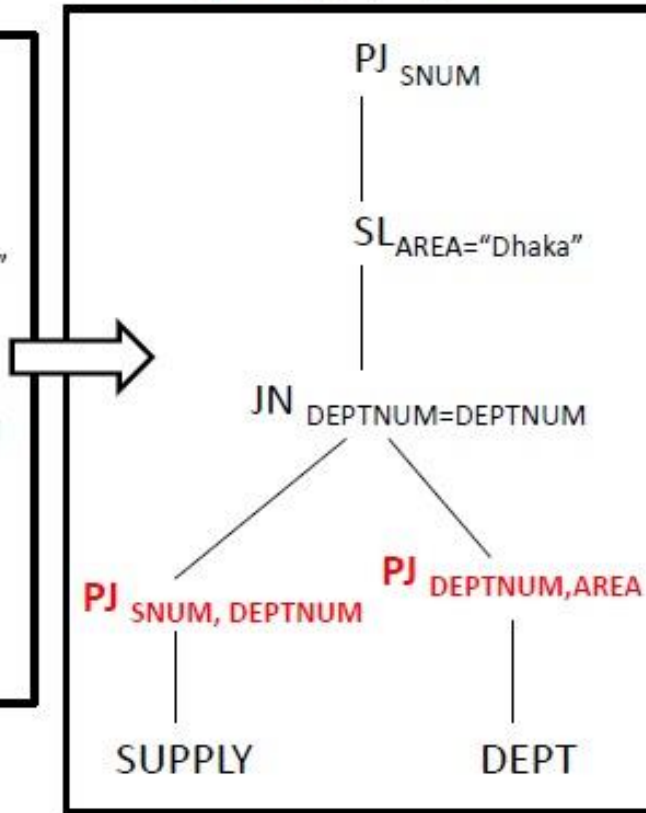
SNUM	S.DEPTNUM	D.DEPTNUM	AREA
11	1	1	Dhaka
41	4	4	Dhaka
71	7	7	Dhaka

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Operator Tree for Q1



After applying Cr-1



1 column, 3 rows !

SNUM
11
41
71

Query before Criterion 1:

PJ_{SNUM} **SL**_{AREA="Dhaka"} (*SUPPLY JN*_{DEPTNUM=DEPTNUM} *DEPT*)

Query After Criterion 1:

PJ_{SNUM} **SL**_{AREA="Dhaka"} (**PJ**_{SNUM,DEPTNUM} *SUPPLY JN*_{DEPTNUM=DEPTNUM}
PJ_{DEPTNUM,AREA} *DEPT*)

They are Equivalent Expressions !

Simplification of Operator Tree

Criterion – 2:

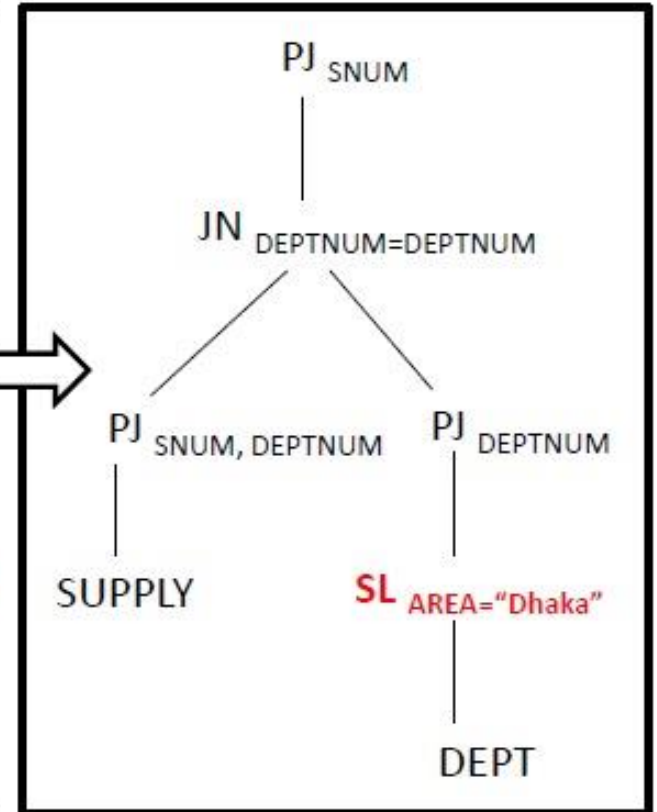
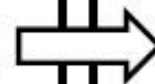
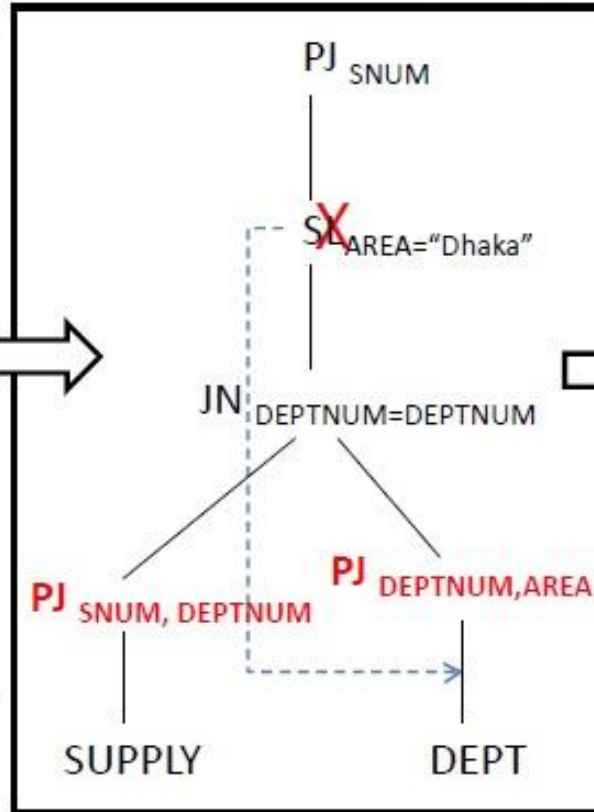
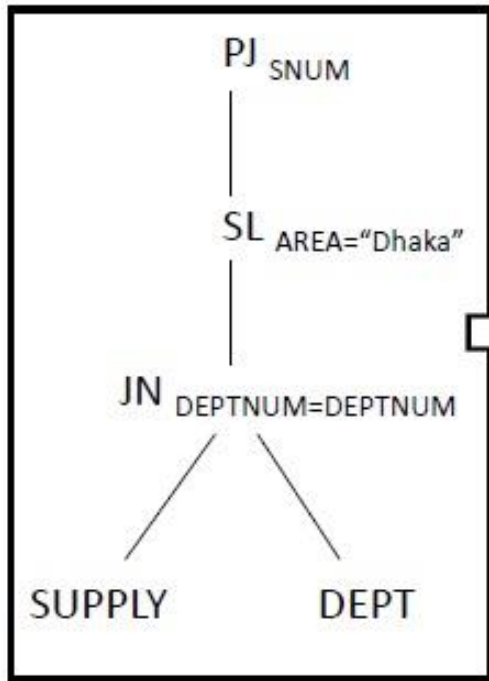
Push *SL* and *PJ* as down as possible in the tree.

- To avoid working on large results (i.e. result of JOIN).

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

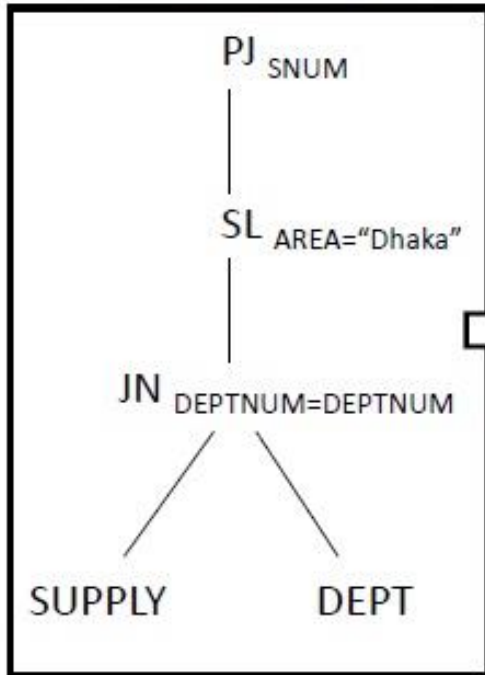
After applying Cr-2

Operator Tree for Q1

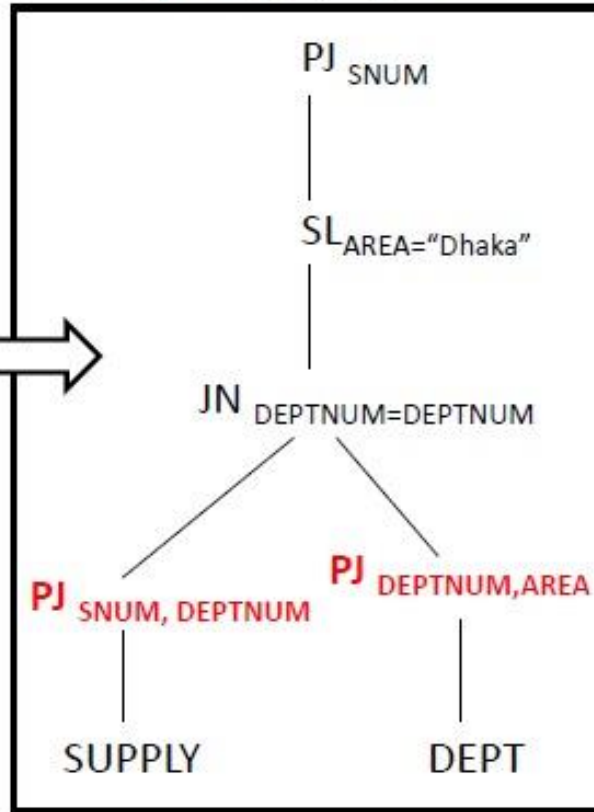


SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Operator Tree for Q1



After applying Cr-1

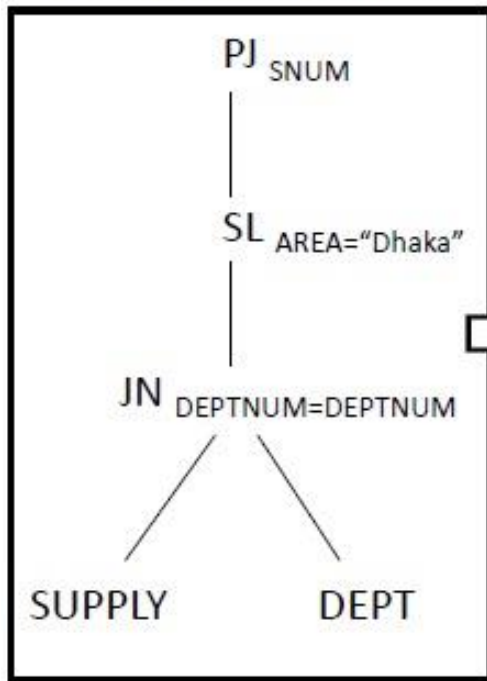


4 columns, 7 rows !

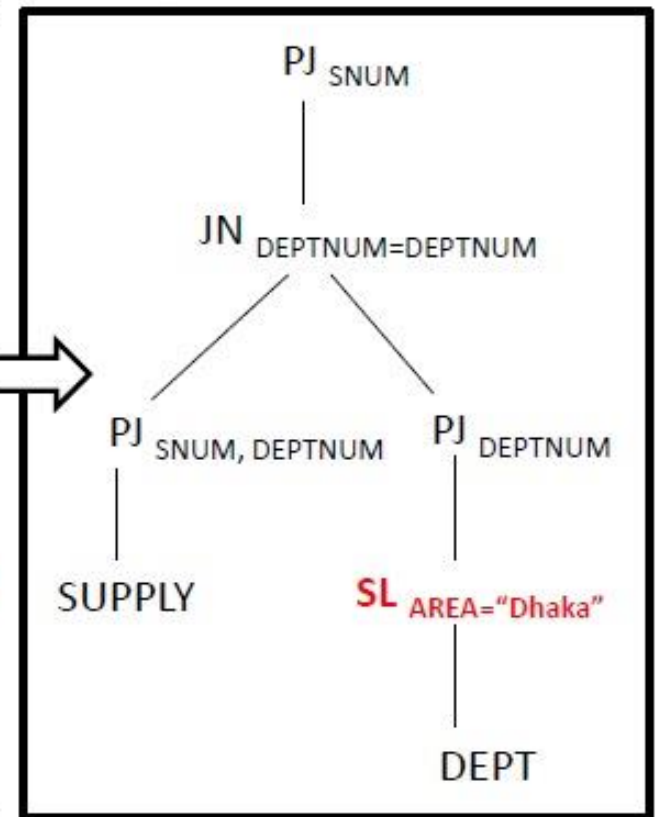
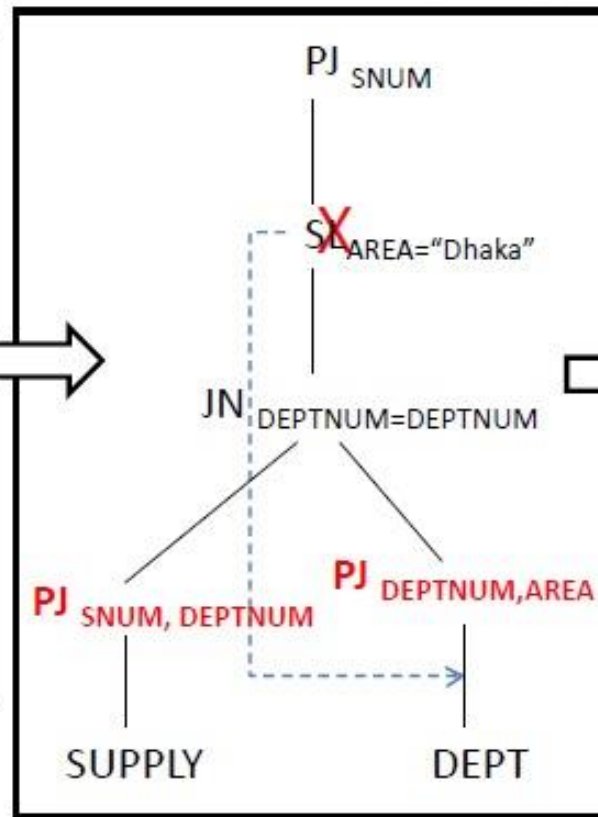
SNUM	S.DEPTNUM	D.DEPTNUM	AREA
11	1	1	Dhaka
12	2	2	Sylhet
13	3	3	Rajshahi
41	4	4	Dhaka
51	5	5	Sylhet
61	6	6	Rajshahi
71	7	7	Dhaka

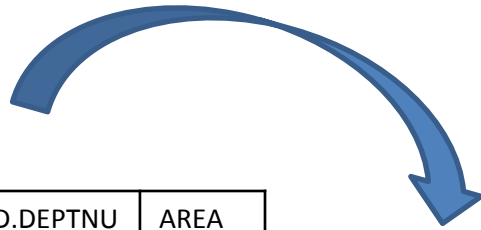
SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Operator Tree for Q1



After applying Cr-2





Only Criteria - 1



Both Criteria 1, 2

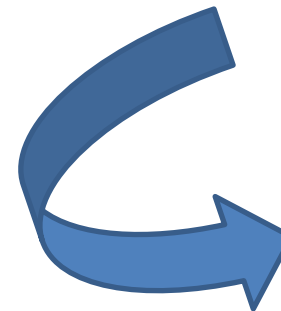
SNUM	S.DEPTNUM	D.DEPTNUM	AREA
11	1	1	Dhaka
12	2	2	Sylhet
13	3	3	Rajshahi
41	4	4	Dhaka
51	5	5	Sylhet
61	6	6	Rajshahi
71	7	7	Dhaka

SNUM	S.DEPTNUM	D.DEPTNUM	AREA
11	1	1	Dhaka
41	4	4	Dhaka
71	7	7	Dhaka

4 columns, 3 rows !

SNUM	S.DEPTNUM	D.DEPTNUM
11	1	1
41	4	4
71	7	7

3 columns, 3 rows !



SNUM
11
41
71

4 columns, 7 rows !

Select Area = Dhaka First, So we do not need the area column anymore

Query before Criterion 1 & 2:

PJ_{SNUM} **SL**_{AREA="Dhaka"} (*SUPPLY JN*_{DEPTNUM=DEPTNUM} *DEPT*)

Query After Criterion 1 & 2:

PJ_{SNUM} (**PJ**_{SNUM,DEPTNUM} *SUPPLY JN*_{DEPTNUM=DEPTNUM} **PJ**_{DEPTNUM} **SL**_{AREA="Dhaka"} *DEPT*)

They are Equivalent Expressions !

Questions:

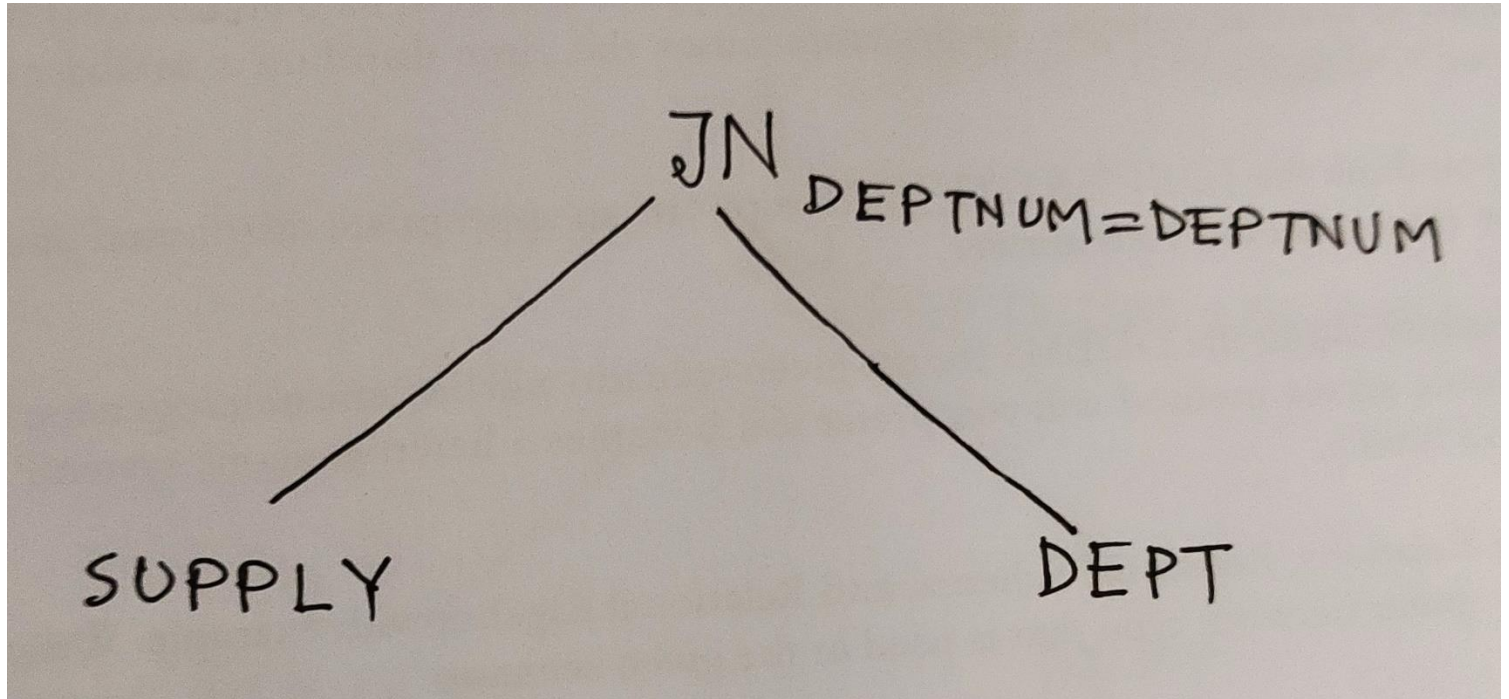
- 1. Should we apply criterion 1 before criterion 2?**
- 2. Can we apply criterion 2 before criterion 1?**
- 3. Can we use either criterion 1 or criterion 2?**
- 4. Should we must have to use criterions?**

Practice

- Try to apply Criterion 1 and/or 2
- Write the equivalent expressions

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

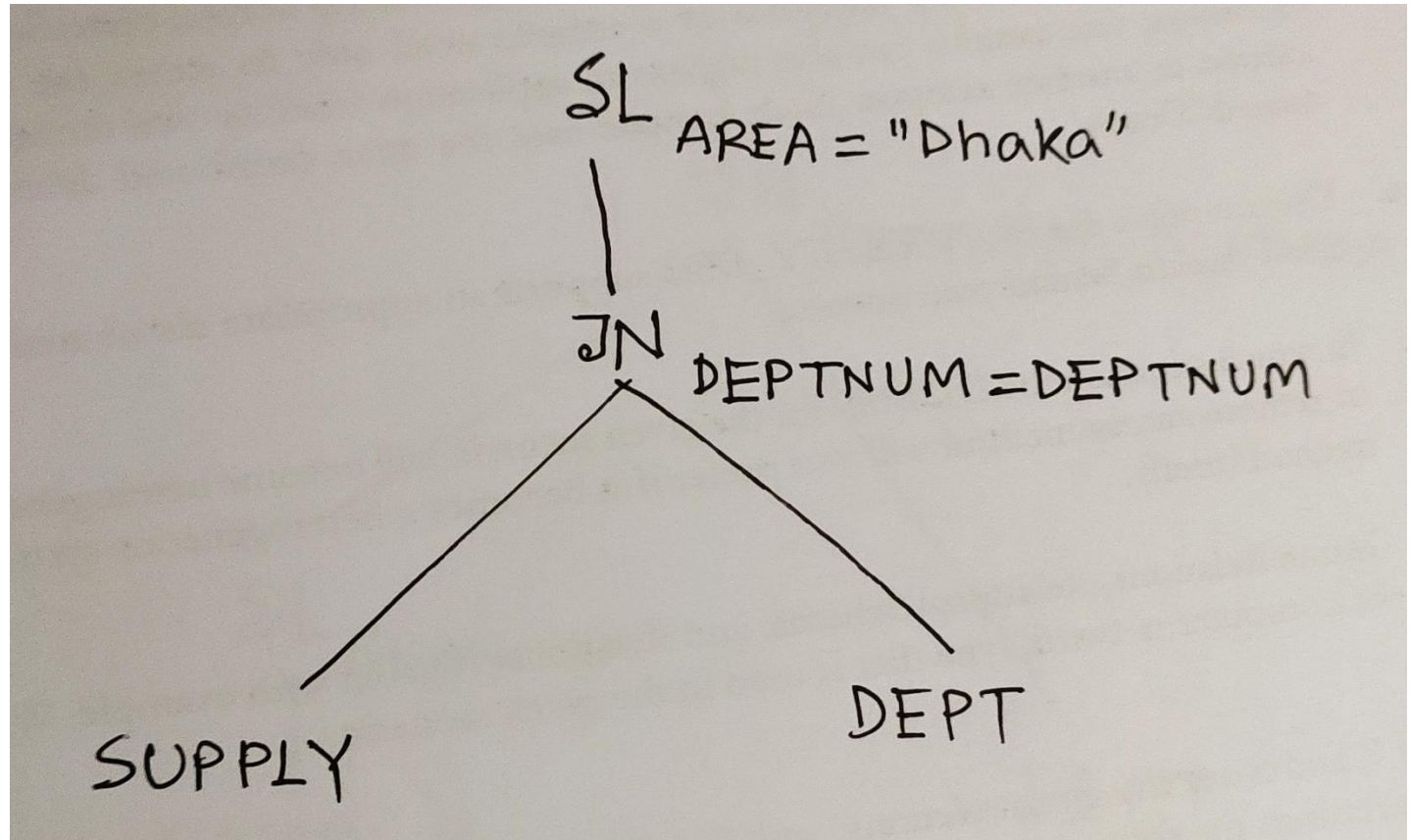
Tree1:



Query: *SUPPLY JN DEPTNUM=DEPTNUM DEPT*

SUPPLY (*SNUM*, *PNUM*, *DEPTNUM*, *QUAN*)
DEPT (*DEPTNUM*, *NAME*, *AREA*, *MGRNUM*)

Tree 2:

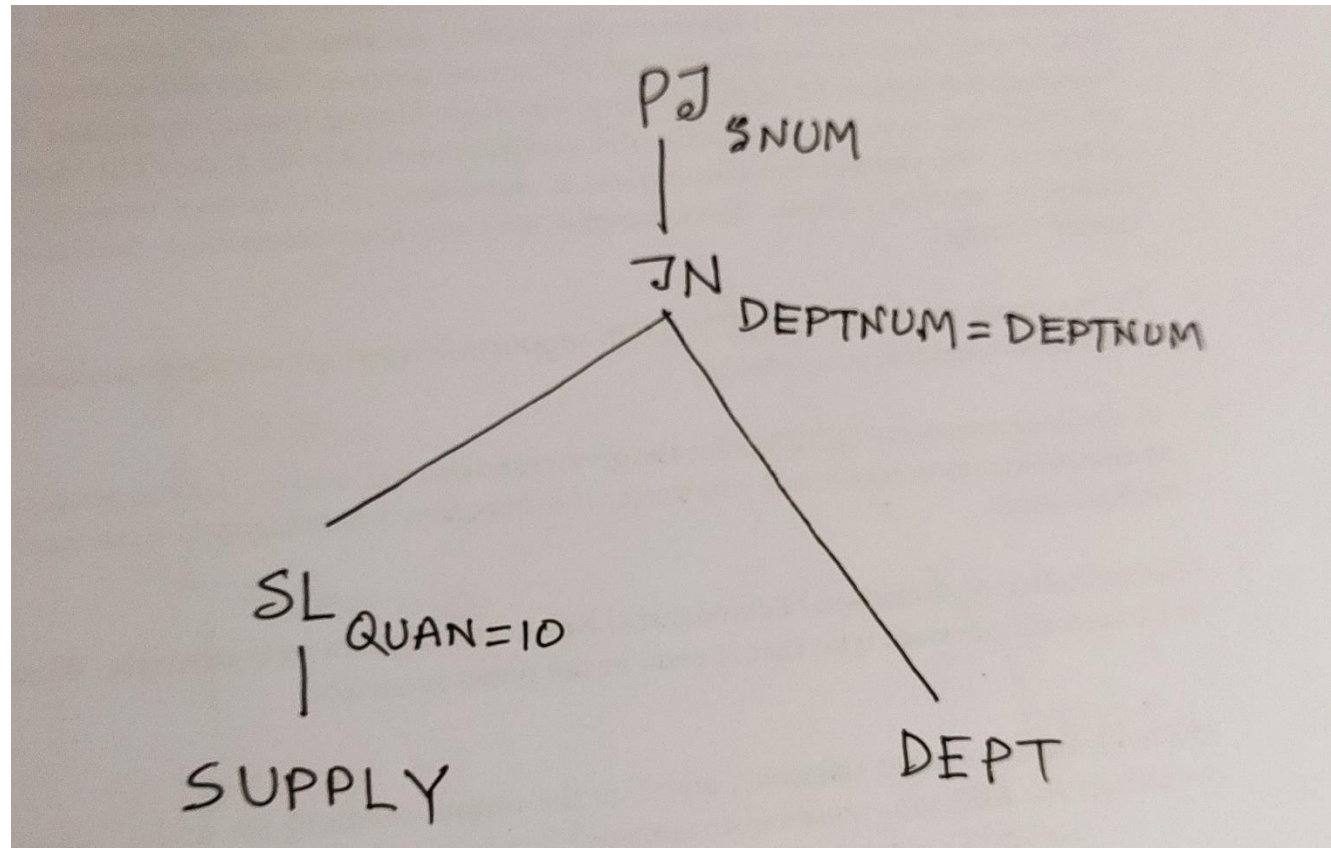


Query: **SL**_{AREA="Dhaka"} (*SUPPLY* **JN**_{DEPTNUM=DEPTNUM} *DEPT*)

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)

DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Tree 3:

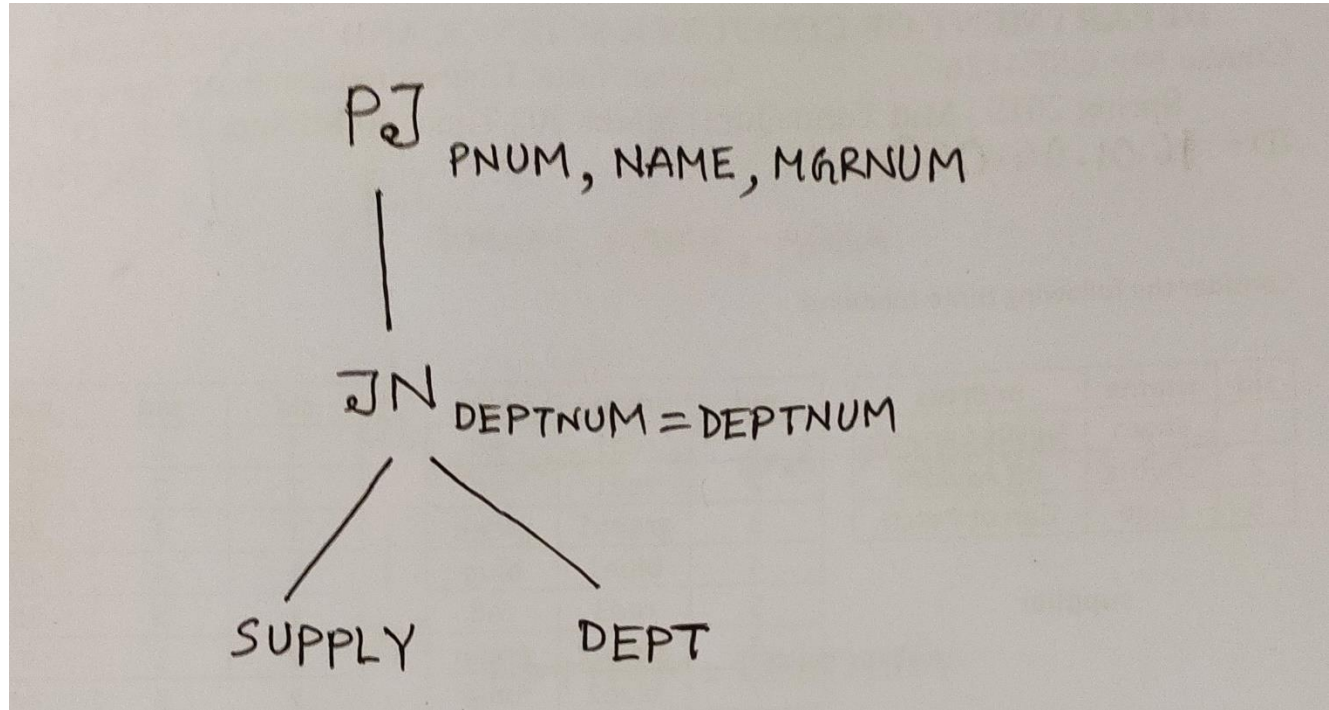


Query: **PJ**_{SNUM} (**SL**_{QUAN=10} *SUPPLY* **JN**_{DEPTNUM=DEPTNUM} *DEPT*)

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)

DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Tree 4:

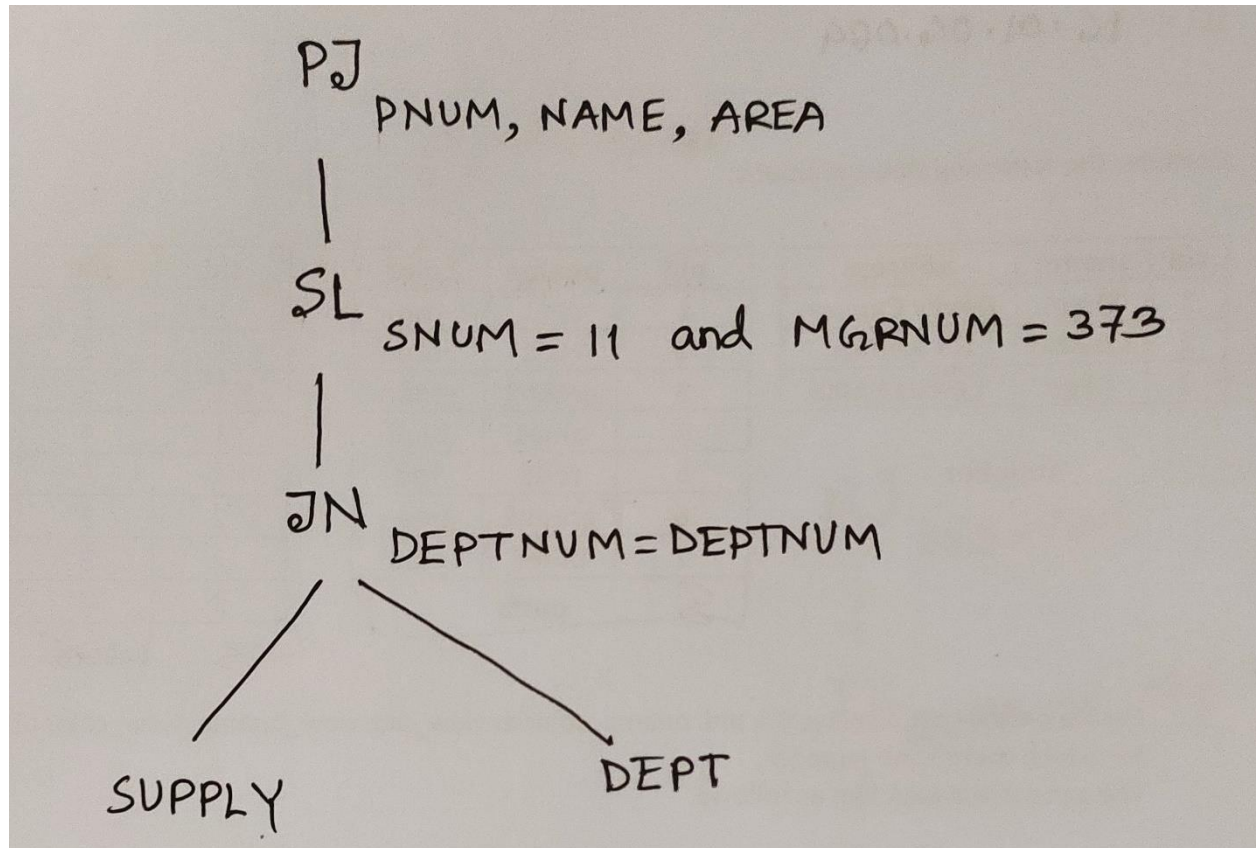


Query: **PJ** *PNUM,NAME,MGRNUM* (*SUPPLY* **JN** *DEPTNUM=DEPTNUM* *DEPT*)

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)

DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Tree 5:

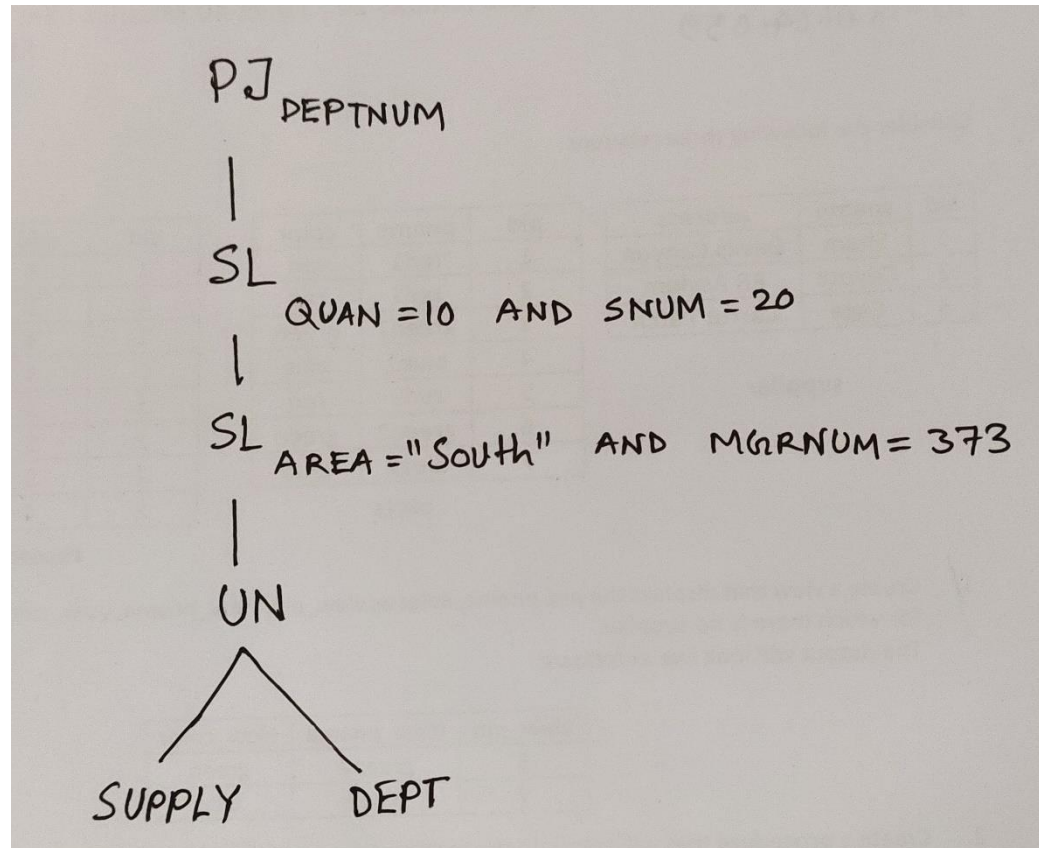


Query: **PJ** *PNUM,NAME,AREA* **SL** *SNUM=11 and MGRNUM=373* (**SUPPLY JN**
DEPTNUM=DEPTNUM **DEPT**)

SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)

DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Tree 6:



Query: **PJ**_{DEPTNUM} **SL**_{QUAN=10 AND SNUM = 20} **SL**_{AREA="South" AND MGRNUM = 373} (**SUPPLY UN DEPT**)

Tree 7:

EMP (EMPNUM, DEPTNUM, NAME, SAL, AGE)
DEPT (DEPTNUM, NAME, AREA, MGRNUM)

Query:

PJ EMP.NAME **SL** SAL <= 35K (*EMP* **JN** DEPTNUM=DEPTNUM **SL** MGRNUM=373 *DEPT*)

1. Draw the operator tree.
2. Apply Criterion 1 and/or 2.