# 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 criterions for simplification of operator tree

# Equivalent Expressions of Queries

- **Q1:** $\Pi_{\text{NAME, DEPTNUM}}\ \sigma_{\text{DEPTNUM} = 15}\ \textit{EMP}$
- **Q2:** $\sigma_{\text{DEPTNUM} = 15}\ \Pi_{\text{NAME, DEPTNUM}}\ \textit{EMP}$

$$Q1 \leftrightarrow Q2$$

# Simplification of Operator Tree

*Criterion – 1*:

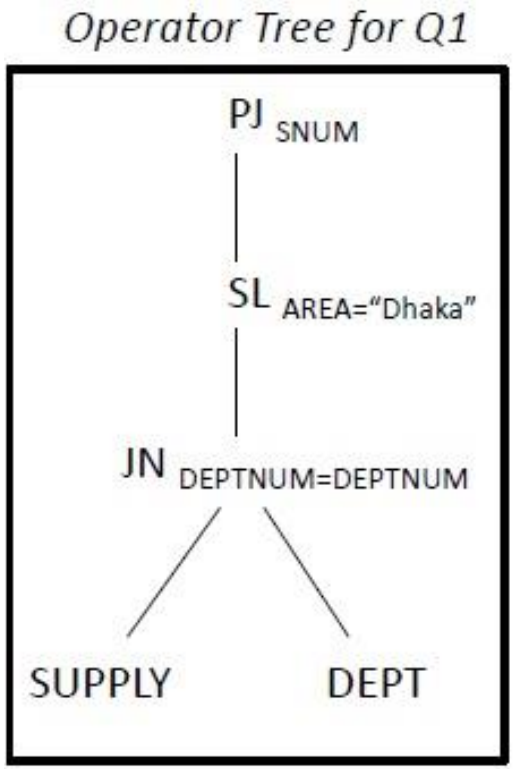
Appropriate introduce of ***SL*** and ***PJ*** in the tree.

- To get rid of unnecessary attributes.

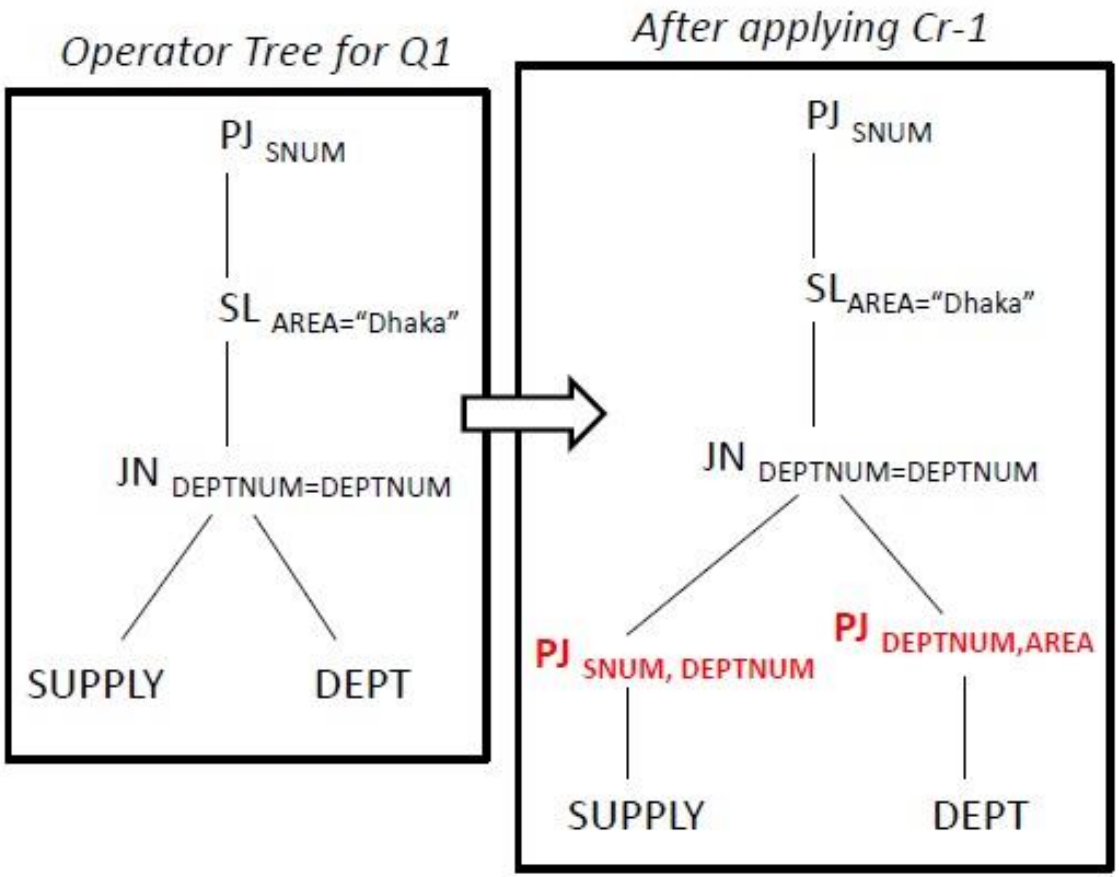
*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*

*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

Q1: **PJ** $_{\text{SNUM}}$ **SL** $_{\text{AREA=“Dhaka”}}$ (*SUPPLY* **JN** $_{\text{DEPTNUM=DEPTNUM}}$ *DEPT*)



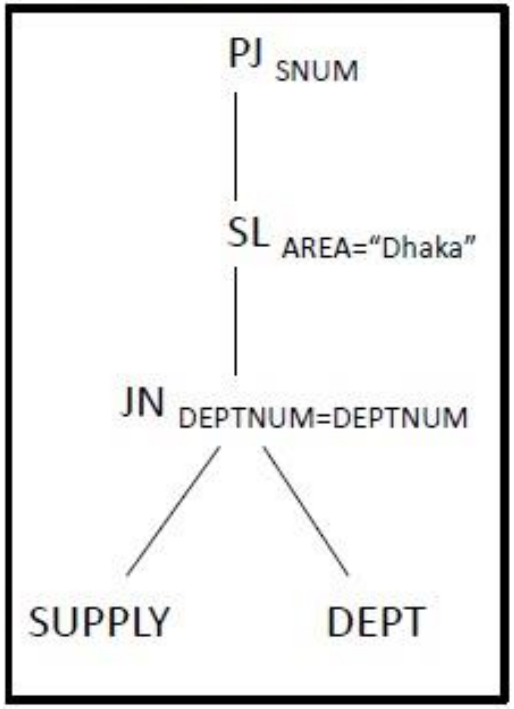
*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*
*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*



*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*

*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

*Operator Tree for Q1*



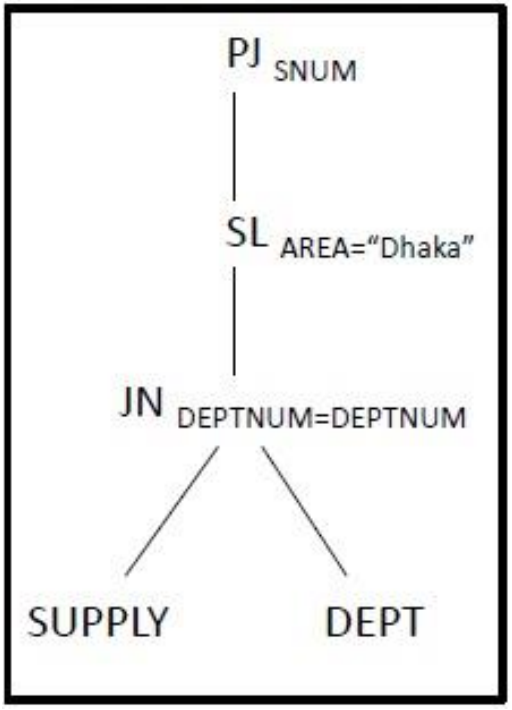

8 columns, 7 rows !

| 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*



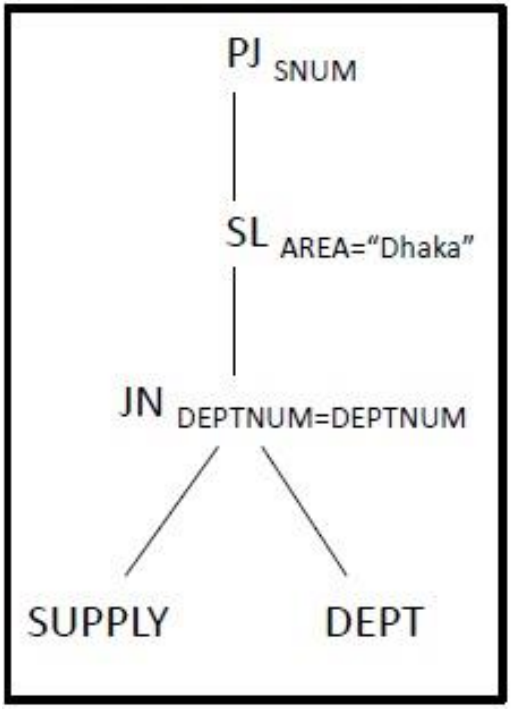

8 columns, 3 rows !

| 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*

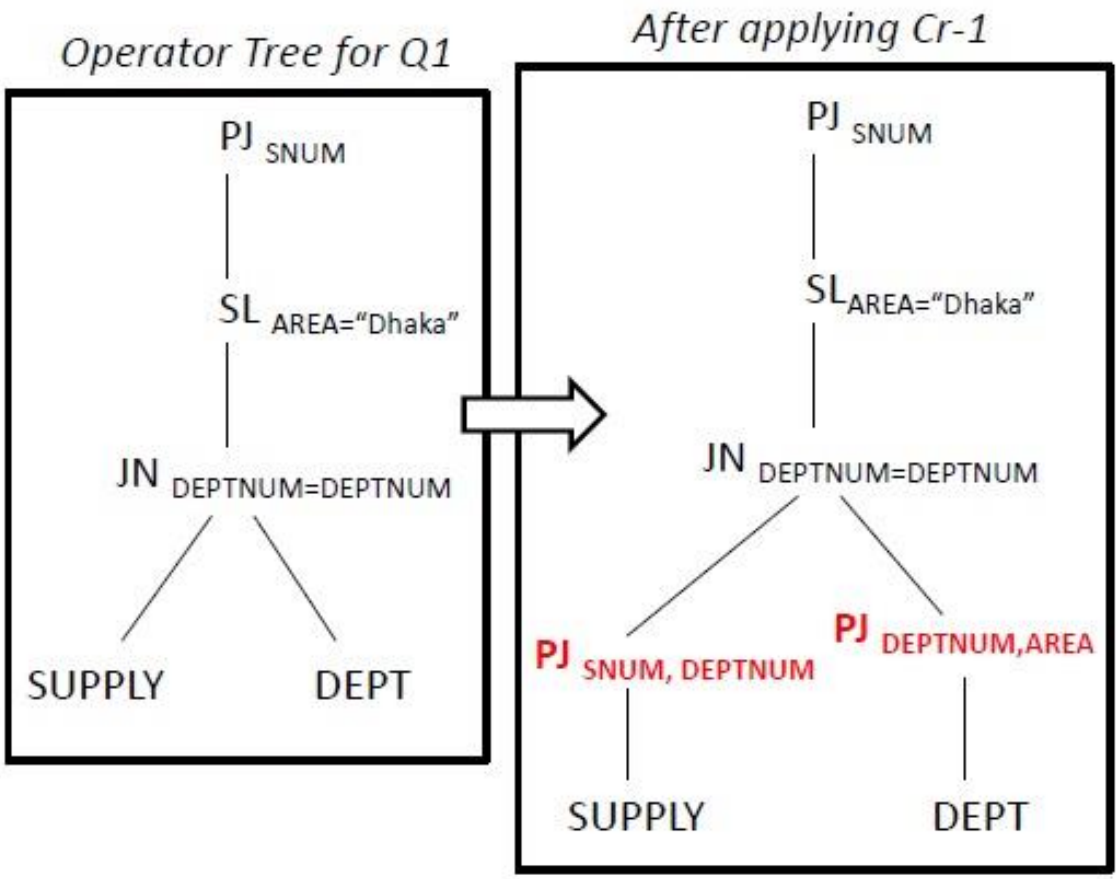



1 column, 3 rows !

| SNUM |
|---|
| 11 |
| 41 |
| 71 |

*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*

*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

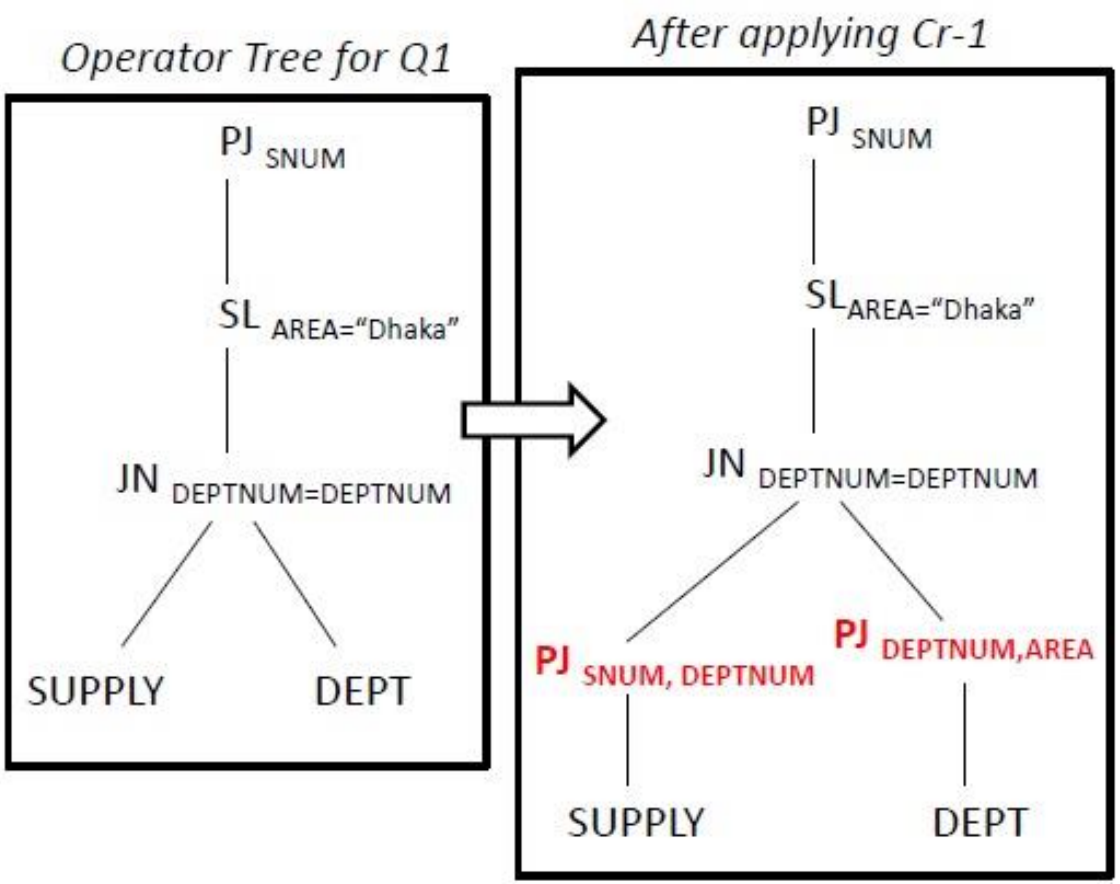



4 columns, 7 rows !

| 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)*

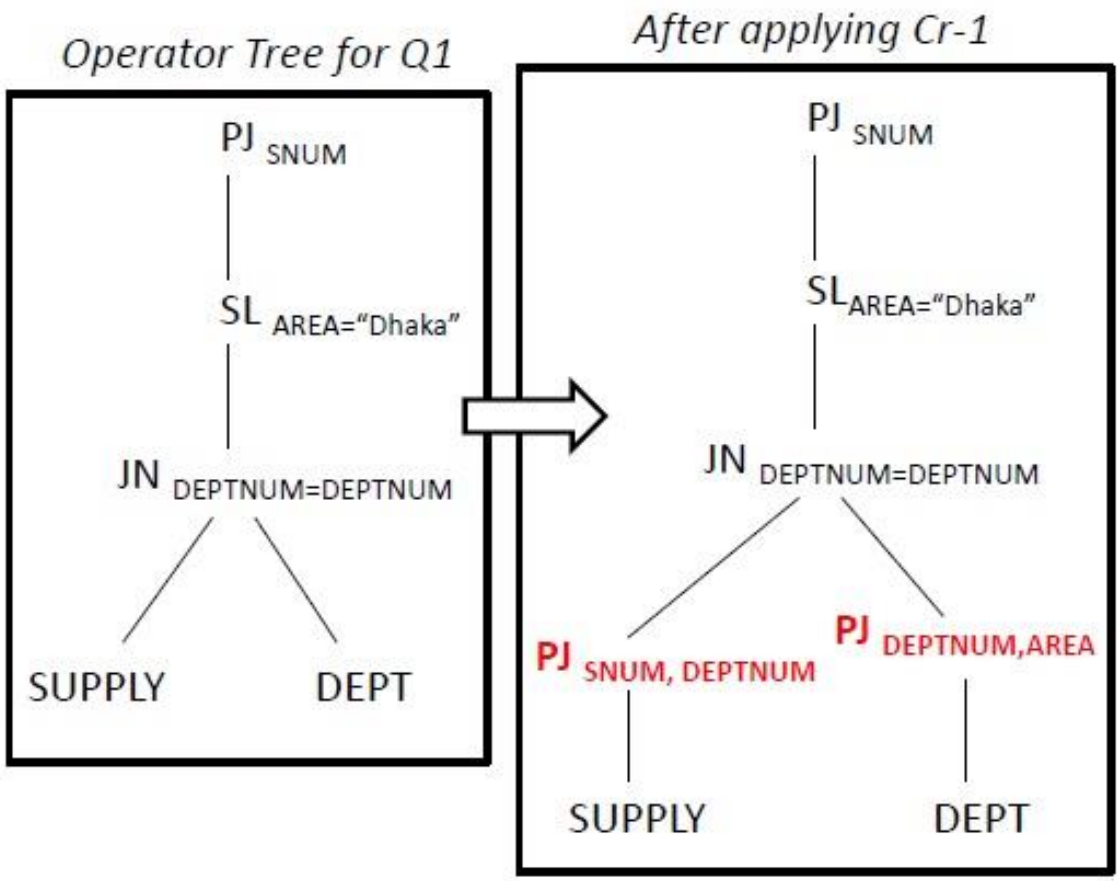



4 columns, 3 rows !

| 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)*




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

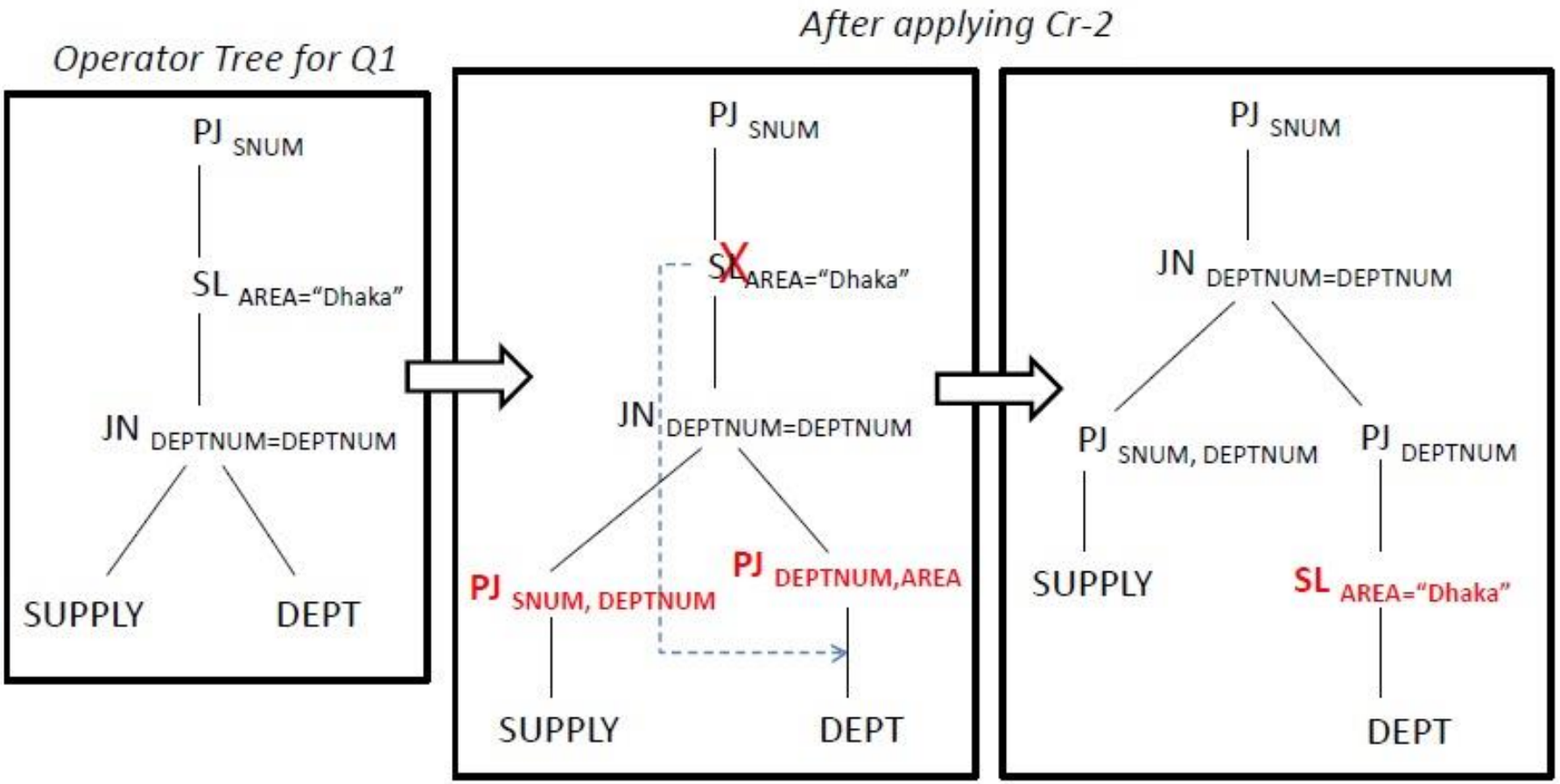
*Criterion – 2:*

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

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

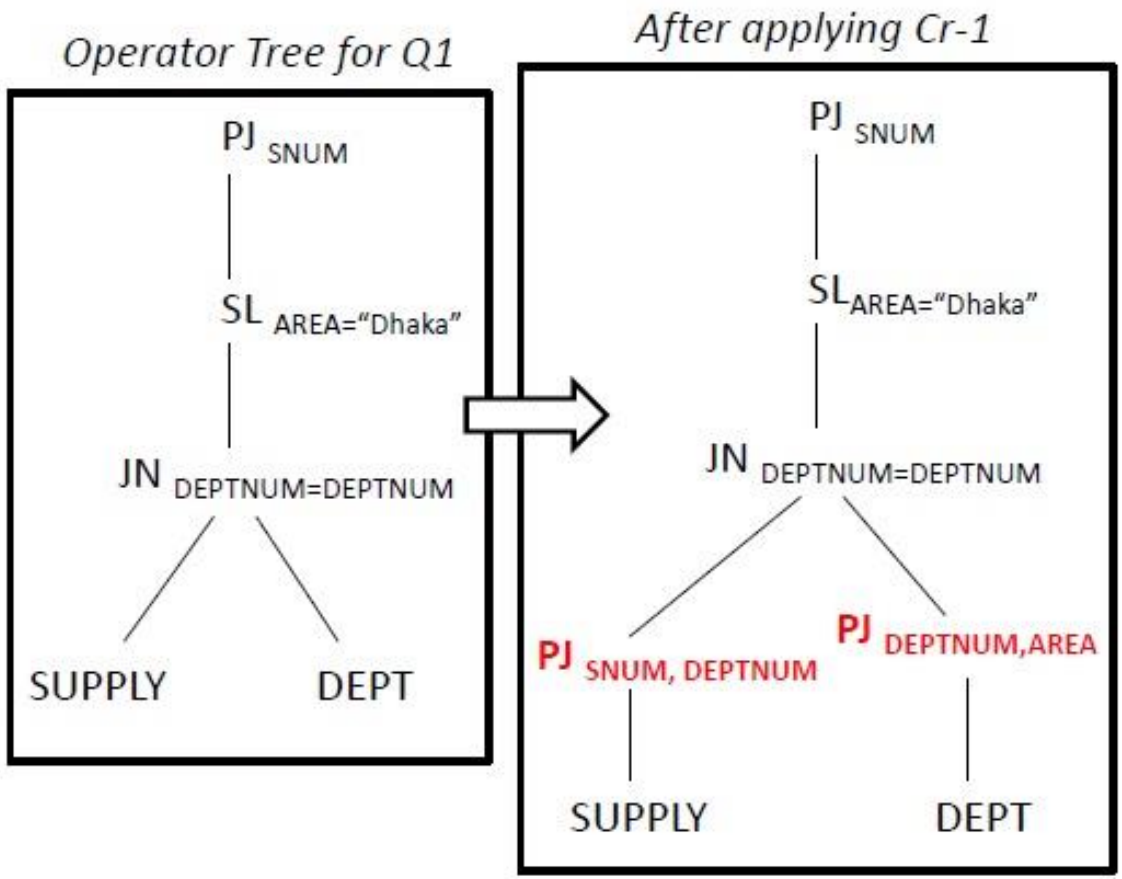
*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*

*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*



*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*

*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

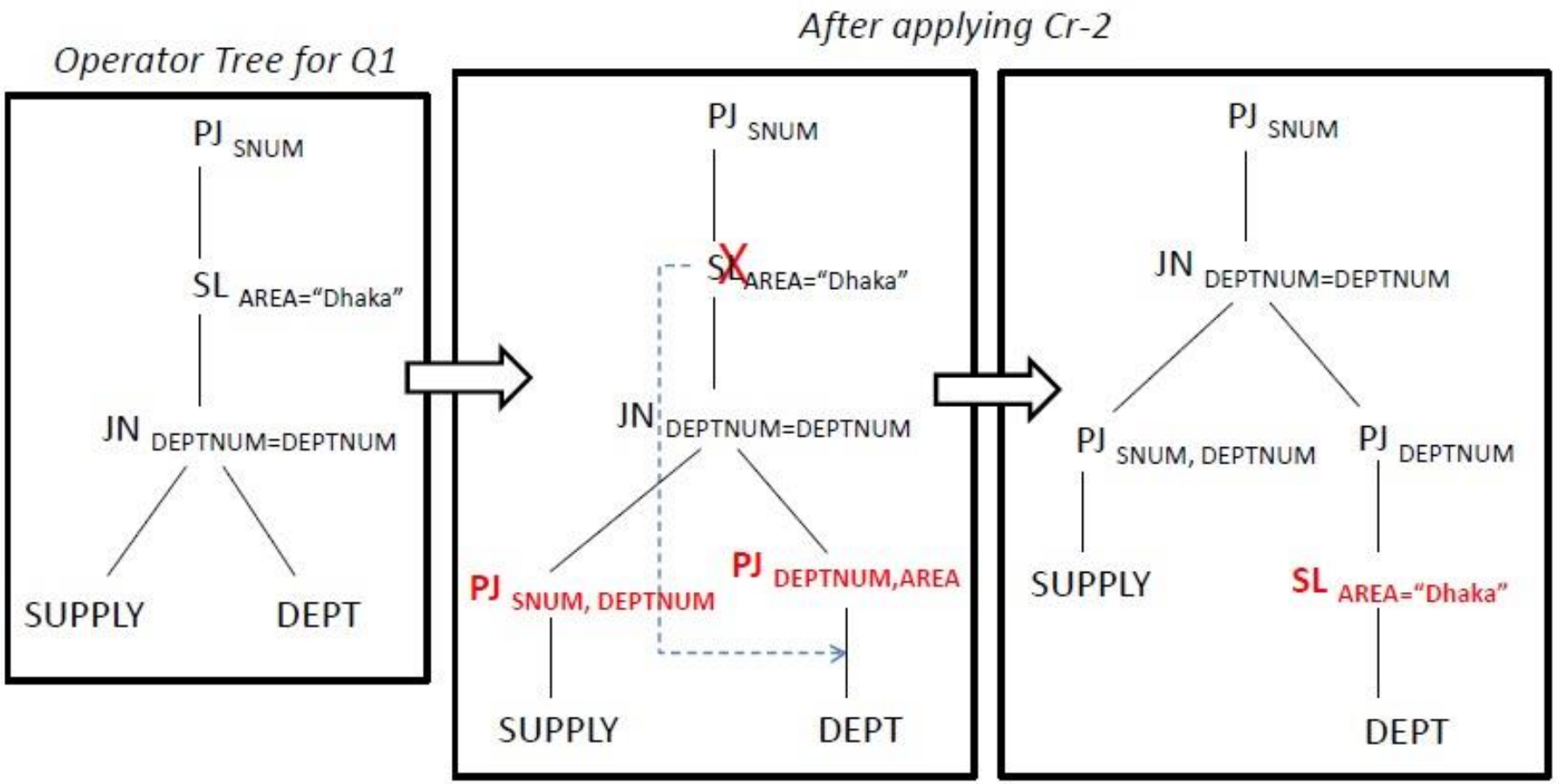



4 columns, 7 rows !

| 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)*



| 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 |

4 columns, 7 rows !

Only Criteria - 1

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

4 columns, 3 rows !

Both Criteria 1, 2

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

3 columns, 3 rows !

| SNUM |
|---|
| 11 |
| 41 |
| 71 |

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:

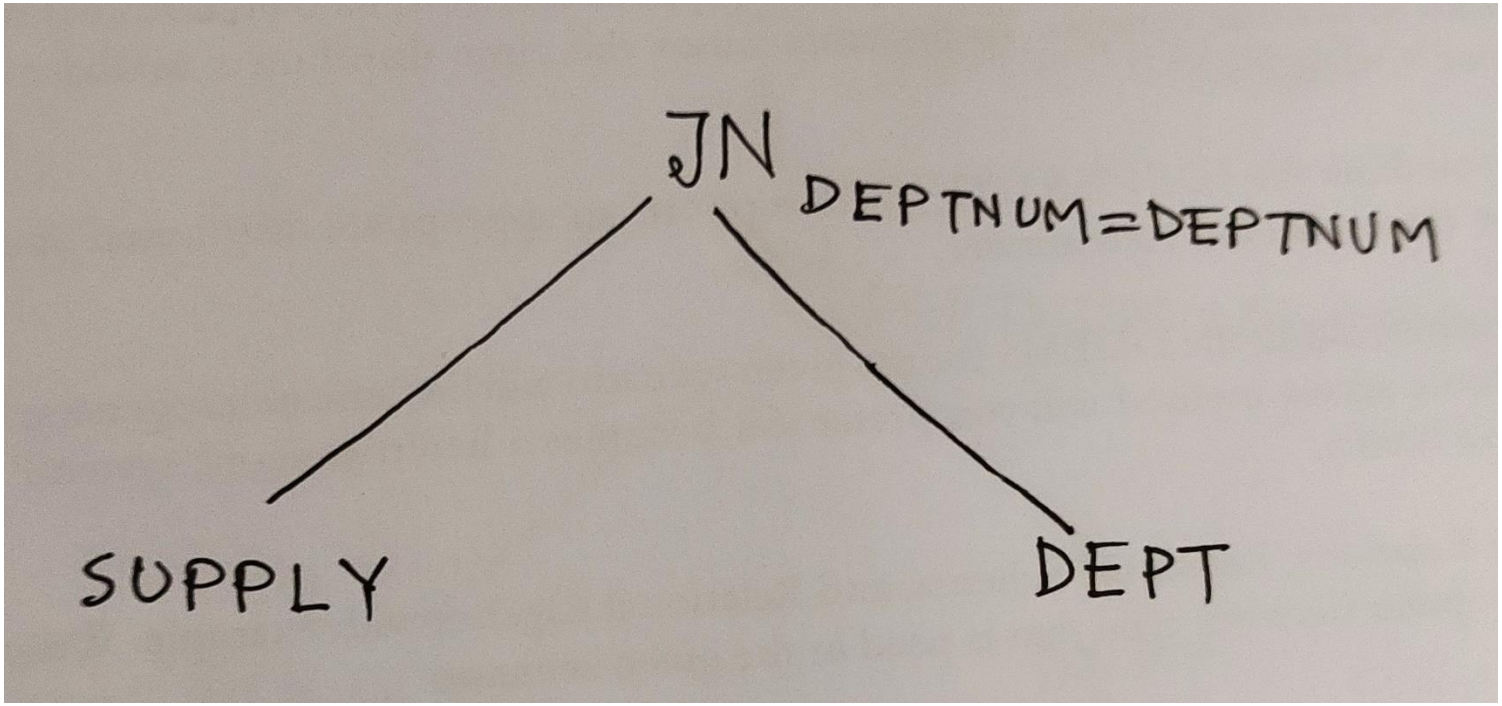
1. **Should we apply criterion 1 before criterion 2?** No
2. **Can we apply criterion 2 before criterion 1?** Yes
3. **Can we use either criterion 1 or criterion 2?** Yes
4. **Should we must have to use criterions?** No

# Practice

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

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

**Tree1:**

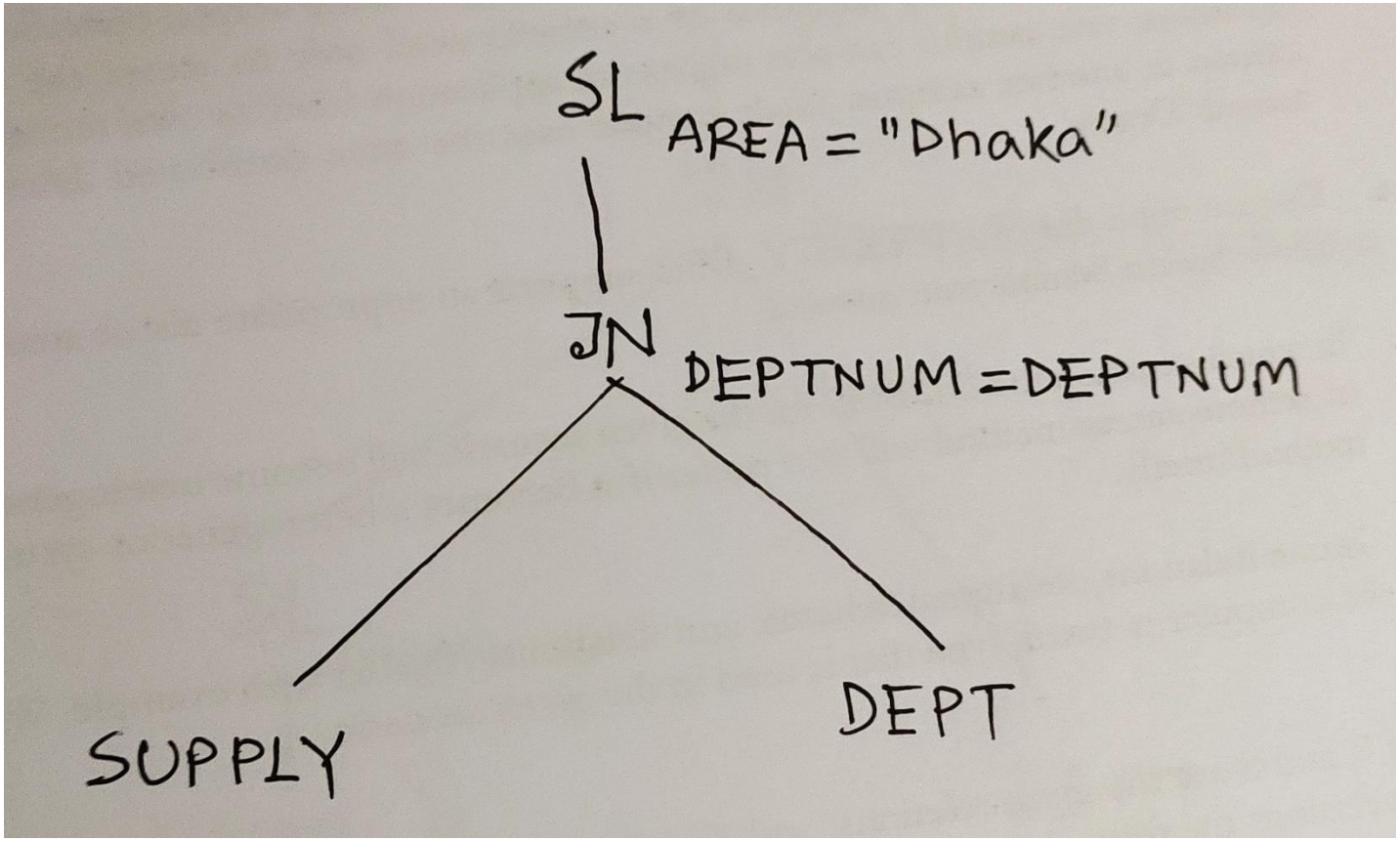



*Query:* *SUPPLY* **JN** $_{\text{DEPTNUM=DEPTNUM}}$ *DEPT*

*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*

*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

**Tree 2:**

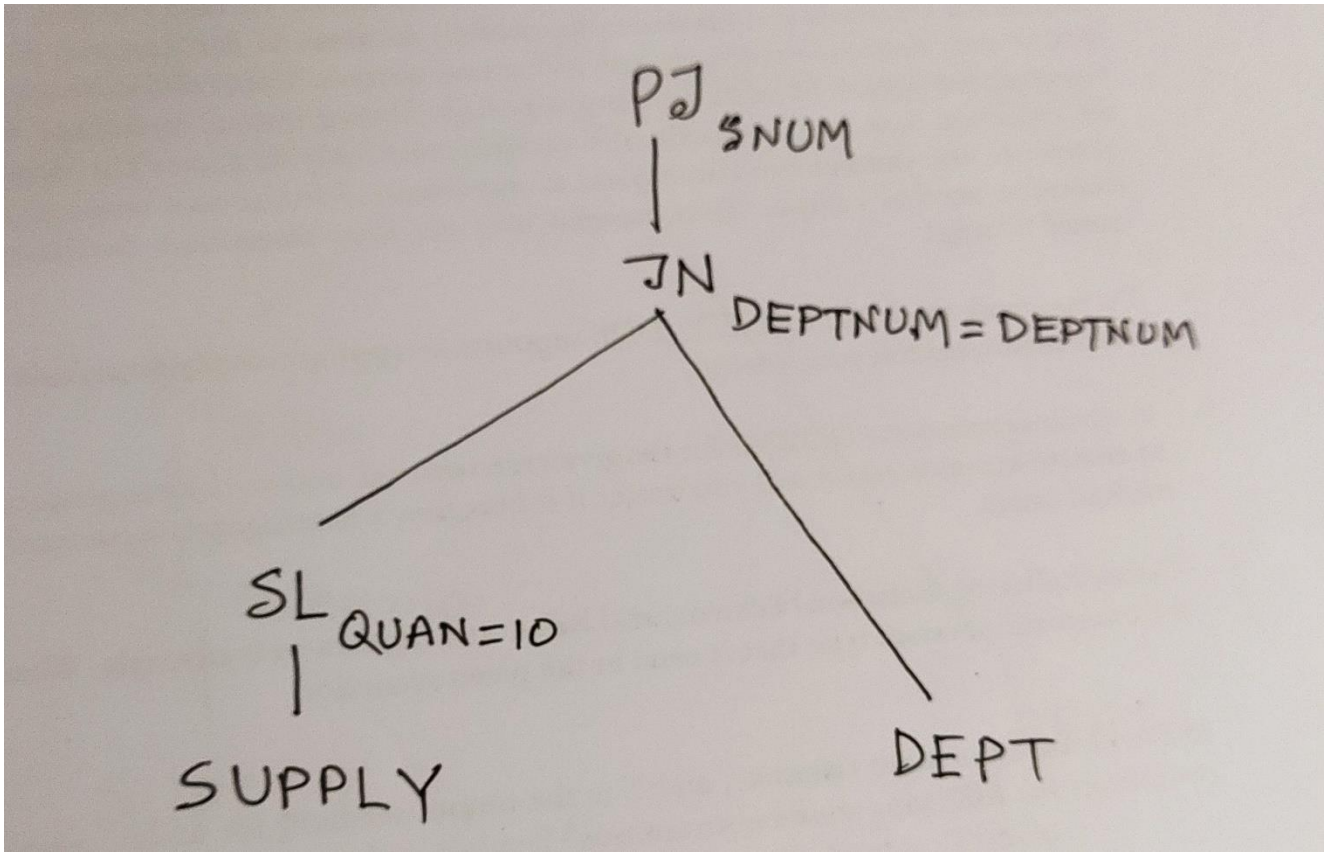



Query: **SL** $_{\text{AREA=“Dhaka”}}$ (*SUPPLY* **JN** $_{\text{DEPTNUM=DEPTNUM}}$ *DEPT*)

*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*

*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

**Tree 3:**

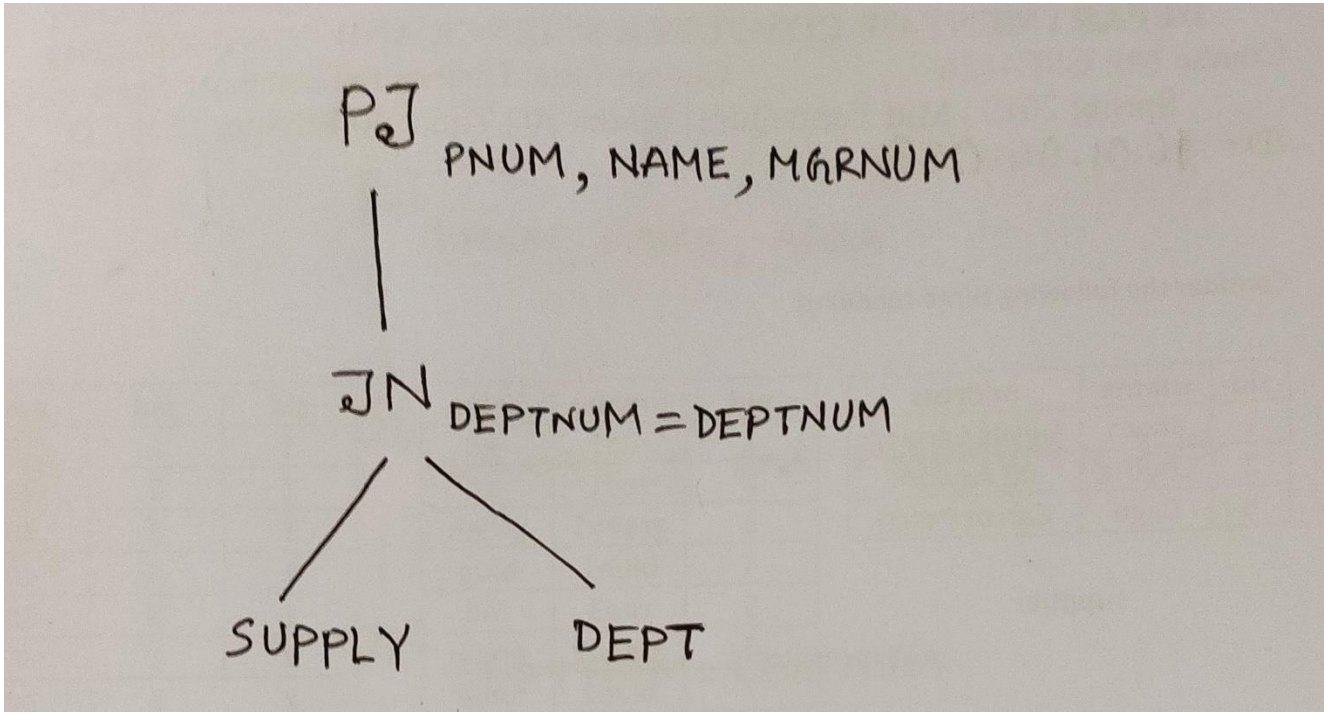



Query: **PJ** $_{\text{SNUM}}$ (**SL** $_{\text{QUAN=10}}$ *SUPPLY* **JN** $_{\text{DEPTNUM=DEPTNUM}}$ *DEPT*)

*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*

*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

**Tree 4:**

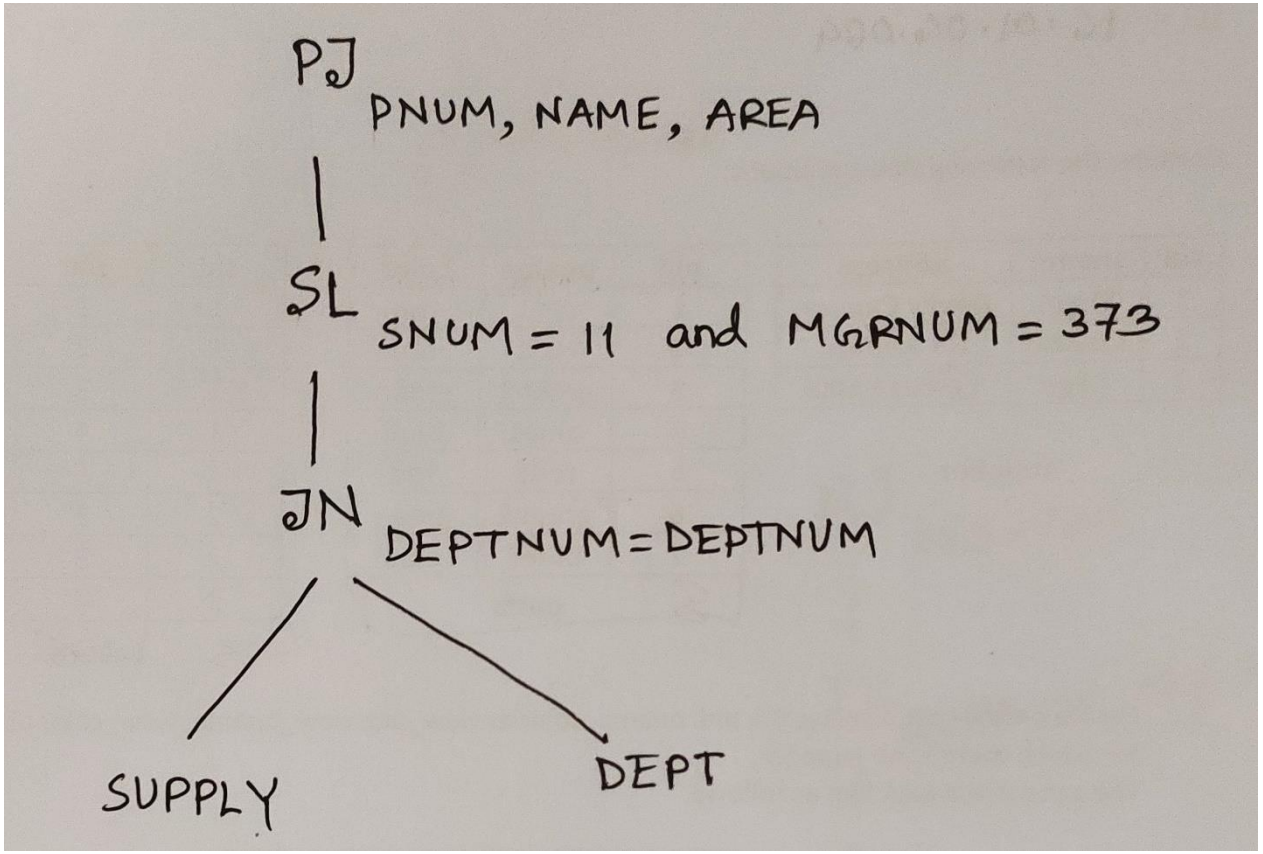



Query: **PJ** PNUM,NAME,MGRNUM (*SUPPLY* **JN** DEPTNUM=DEPTNUM *DEPT*)

*SUPPLY (SNUM, PNUM, DEPTNUM, QUAN)*

*DEPT (DEPTNUM, NAME, AREA, MGRNUM)*

**Tree 5:**

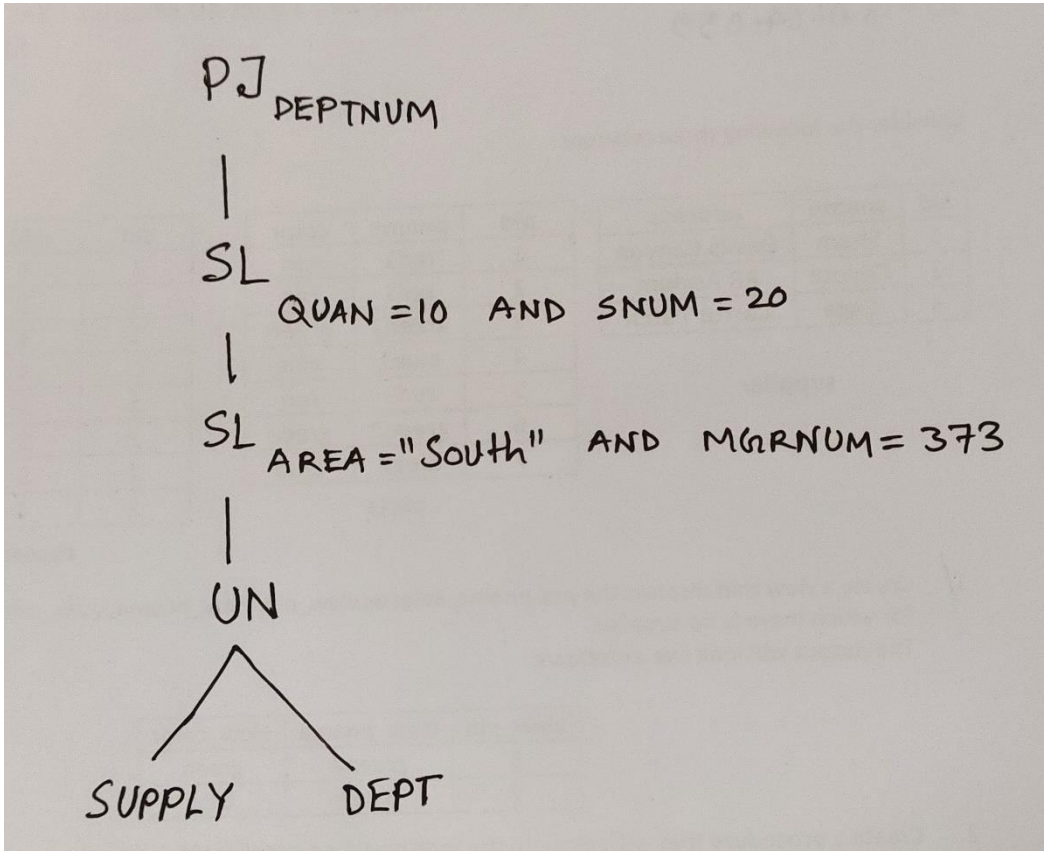



Query: **PJ** PNUM,NAME,AREA **SL** SNUM=11 and MGRNUM=373 (*SUPPLY* **JN** DEPTNUM=DEPTNUM *DEPT*)

*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.